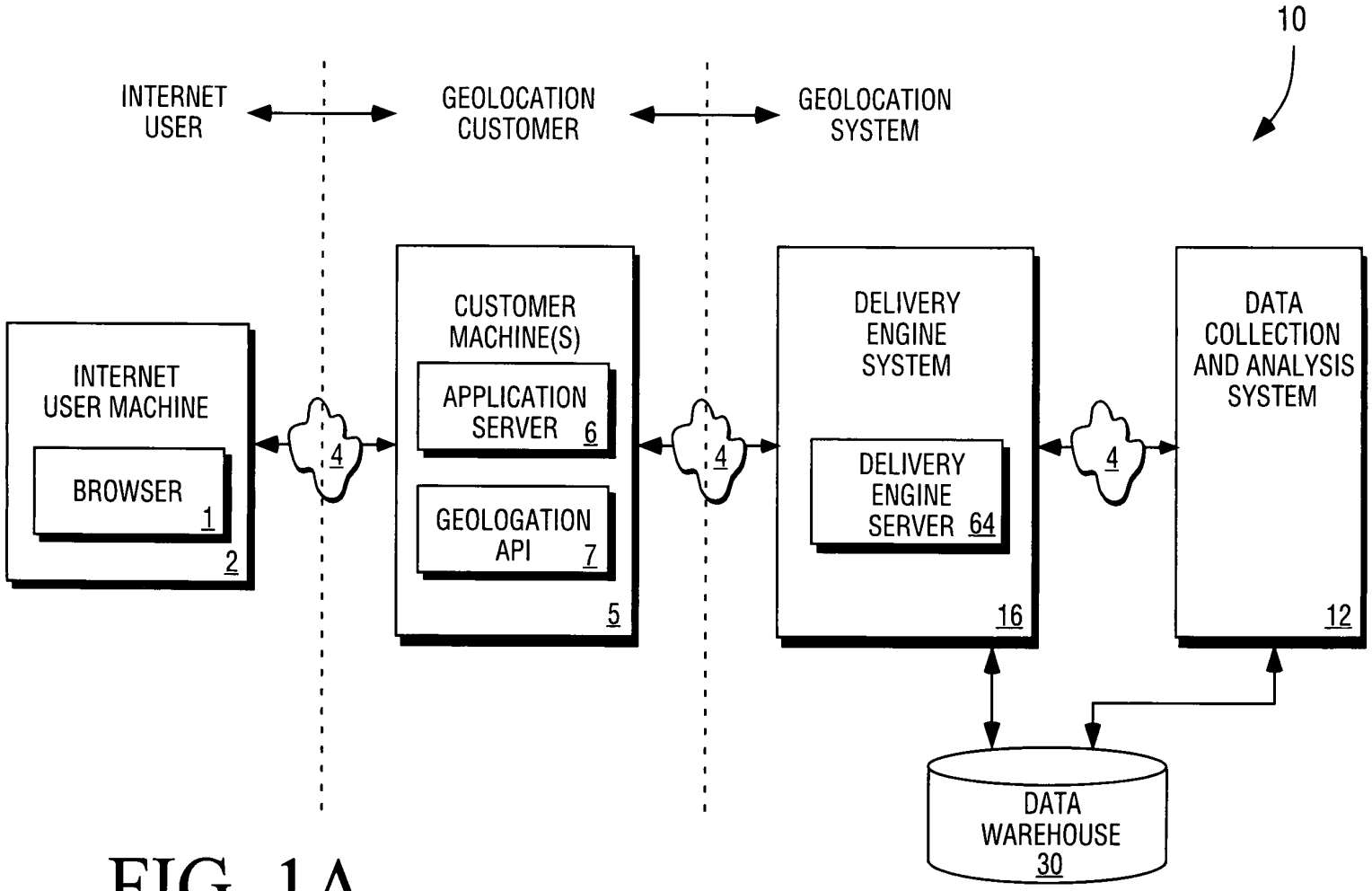

10
INTERNET USER
GEOLOCATION CUSTOMER
GEOLOCATION SYSTEM
INTERNET USER MACHINE
BROWSER 1
2
4
CUSTOMER MACHINE(S)
APPLICATION SERVER 6
GEOLOGATION API 7
5
4
DELIVERY ENGINE SYSTEM
DELIVERY ENGINE SERVER 64
16
4
DATA COLLECTION AND ANALYSIS SYSTEM
12
DATA WAREHOUSE 30

FIG. 1A

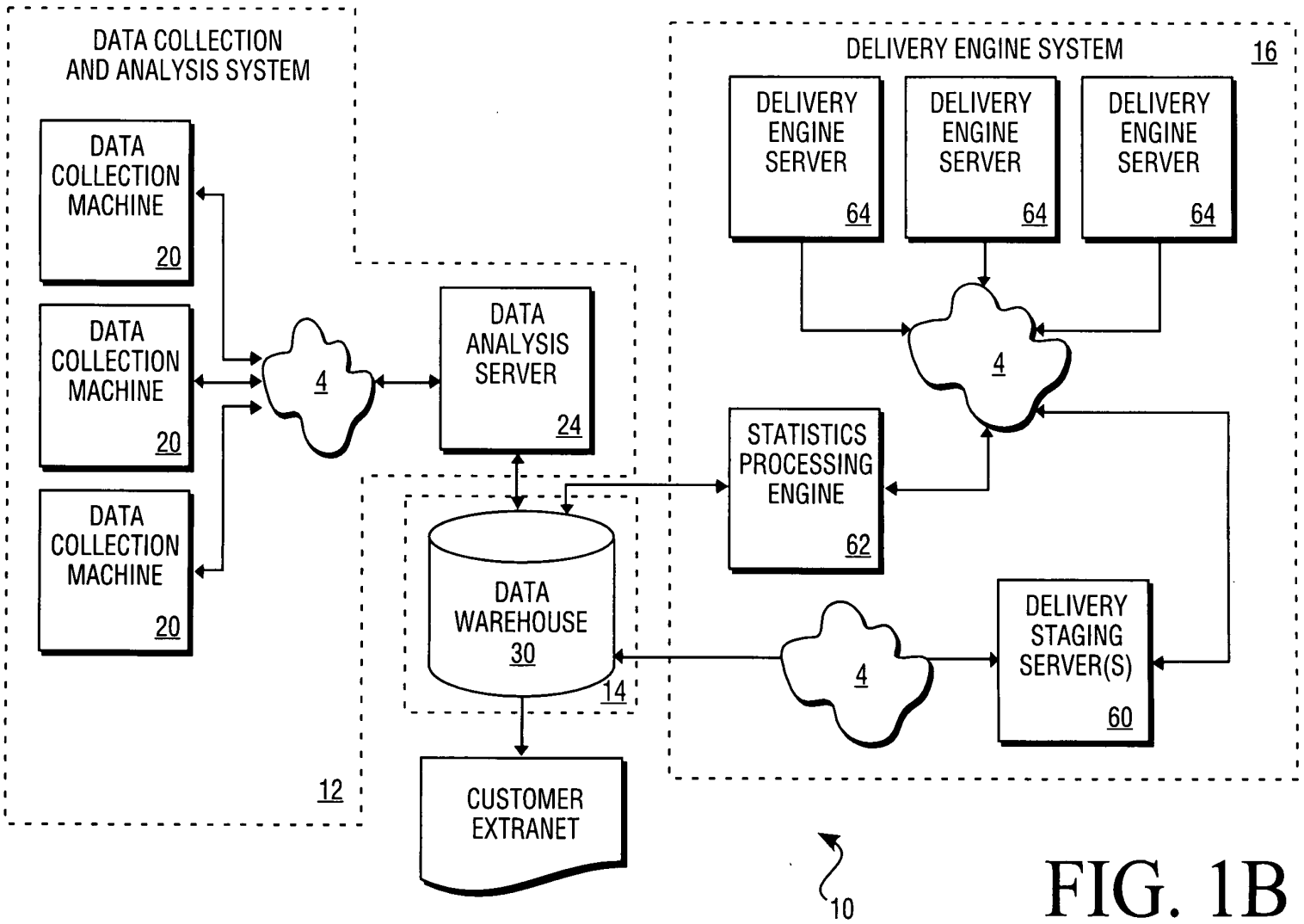

DATA COLLECTION AND ANALYSIS SYSTEM
DATA COLLECTION MACHINE 20
DATA COLLECTION MACHINE 20
DATA COLLECTION MACHINE 20
4
DATA ANALYSIS SERVER 24
DATA WAREHOUSE 30
14
12
CUSTOMER EXTRANET
DELIVERY ENGINE SYSTEM 16
DELIVERY ENGINE SERVER 64
DELIVERY ENGINE SERVER 64
DELIVERY ENGINE SERVER 64
4
STATISTICS PROCESSING ENGINE 62
4
DELIVERY STAGING SERVER(S) 60
10

FIG. 1B

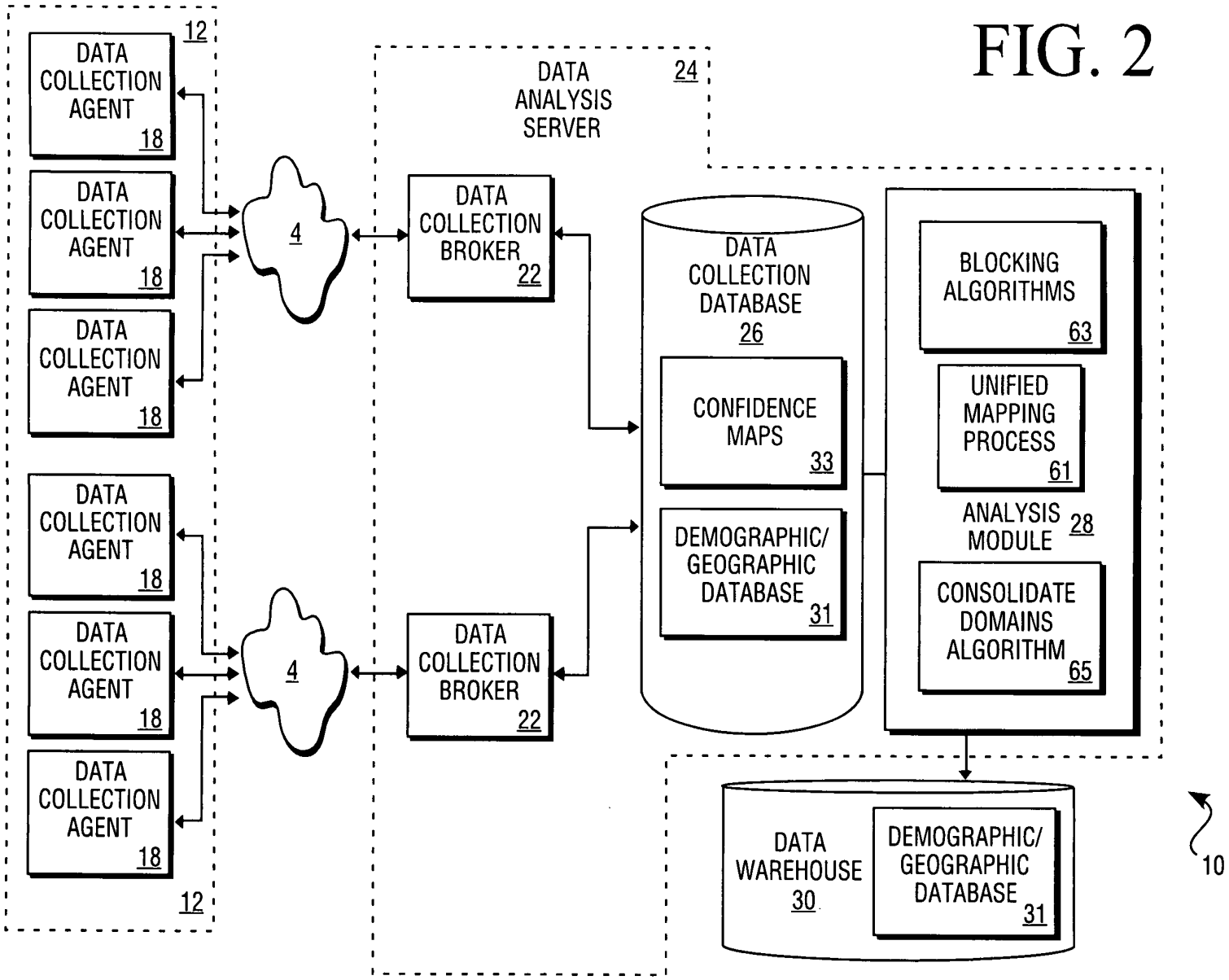

12
DATA COLLECTION AGENT 18
DATA COLLECTION AGENT 18
DATA COLLECTION AGENT 18
DATA COLLECTION AGENT 18
DATA COLLECTION AGENT 18
DATA COLLECTION AGENT 18
12
4
4
DATA ANALYSIS SERVER 24
DATA COLLECTION BROKER 22
DATA COLLECTION BROKER 22
DATA COLLECTION DATABASE 26
CONFIDENCE MAPS 33
DEMOGRAPHIC/ GEOGRAPHIC DATABASE 31
BLOCKING ALGORITHMS 63
UNIFIED MAPPING PROCESS 61
ANALYSIS MODULE 28
CONSOLIDATE DOMAINS ALGORITHM 65
DATA WAREHOUSE 30
DEMOGRAPHIC/ GEOGRAPHIC DATABASE 31
10

FIG. 2

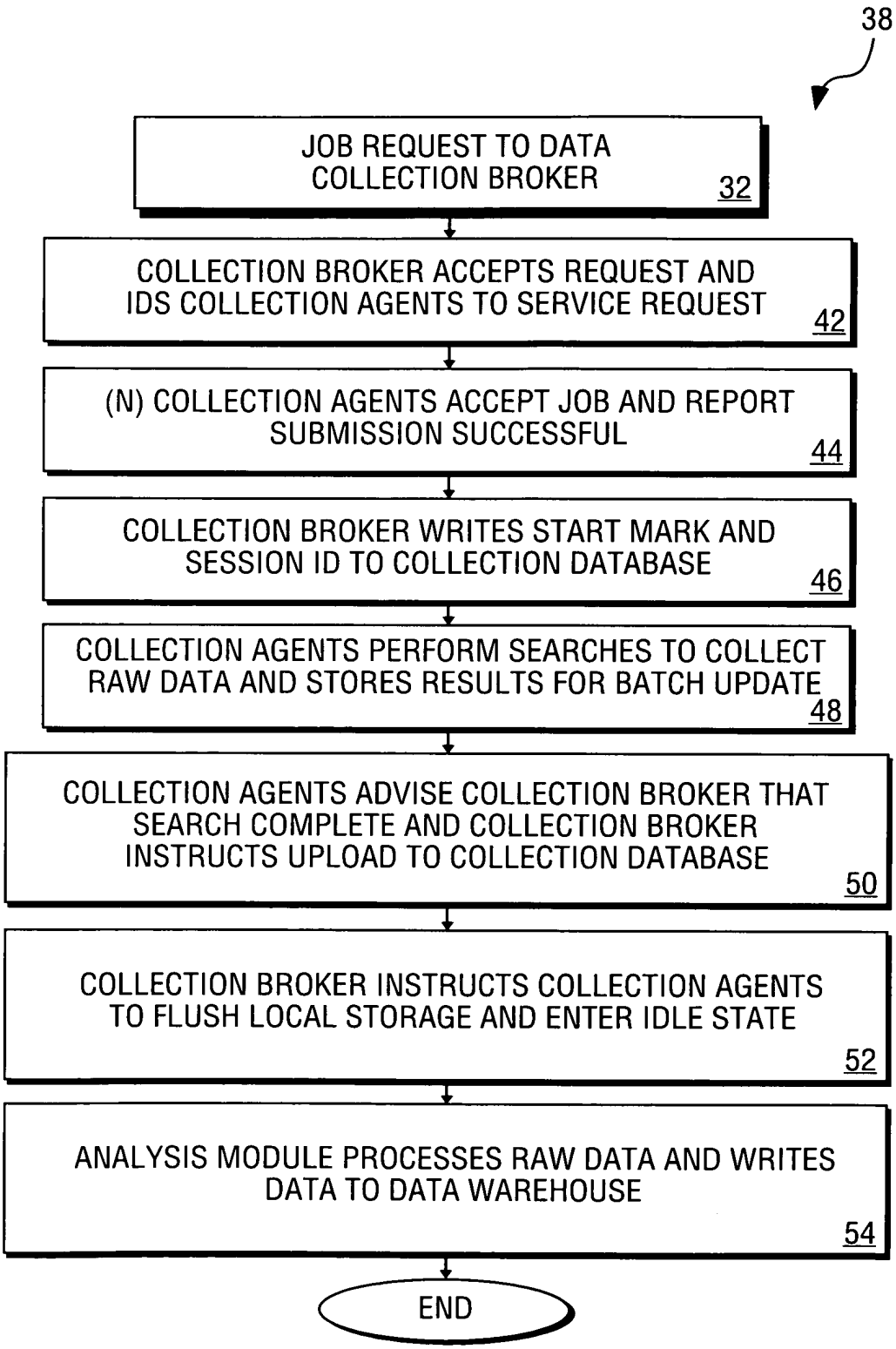

38
JOB REQUEST TO DATA COLLECTION BROKER
32
COLLECTION BROKER ACCEPTS REQUEST AND IDS COLLECTION AGENTS TO SERVICE REQUEST
42
(N) COLLECTION AGENTS ACCEPT JOB AND REPORT SUBMISSION SUCCESSFUL
44
COLLECTION BROKER WRITES START MARK AND SESSION ID TO COLLECTION DATABASE
46
COLLECTION AGENTS PERFORM SEARCHES TO COLLECT RAW DATA AND STORES RESULTS FOR BATCH UPDATE
48
COLLECTION AGENTS ADVISE COLLECTION BROKER THAT SEARCH COMPLETE AND COLLECTION BROKER INSTRUCTS UPLOAD TO COLLECTION DATABASE
50
COLLECTION BROKER INSTRUCTS COLLECTION AGENTS TO FLUSH LOCAL STORAGE AND ENTER IDLE STATE
52
ANALYSIS MODULE PROCESSES RAW DATA AND WRITES DATA TO DATA WAREHOUSE
54
END

FIG. 3

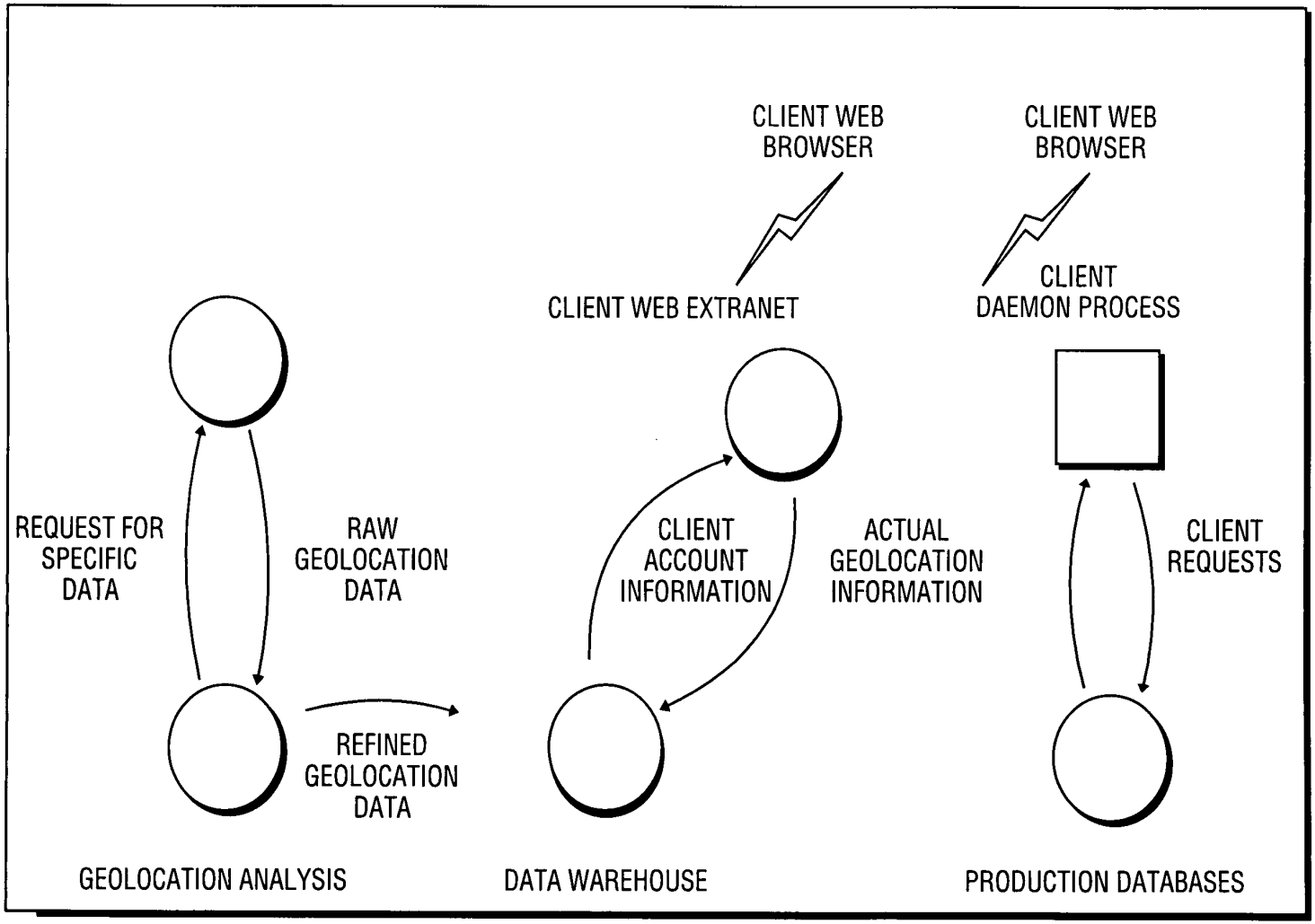

CLIENT WEB BROWSER
CLIENT WEB BROWSER
CLIENT WEB EXTRANET
CLIENT DAEMON PROCESS
REQUEST FOR SPECIFIC DATA
RAW GEOLOCATION DATA
CLIENT ACCOUNT INFORMATION
ACTUAL GEOLOCATION INFORMATION
CLIENT REQUESTS
REFINED GEOLOCATION DATA
GEOLOCATION ANALYSIS
DATA WAREHOUSE
PRODUCTION DATABASES

FIG. 4A

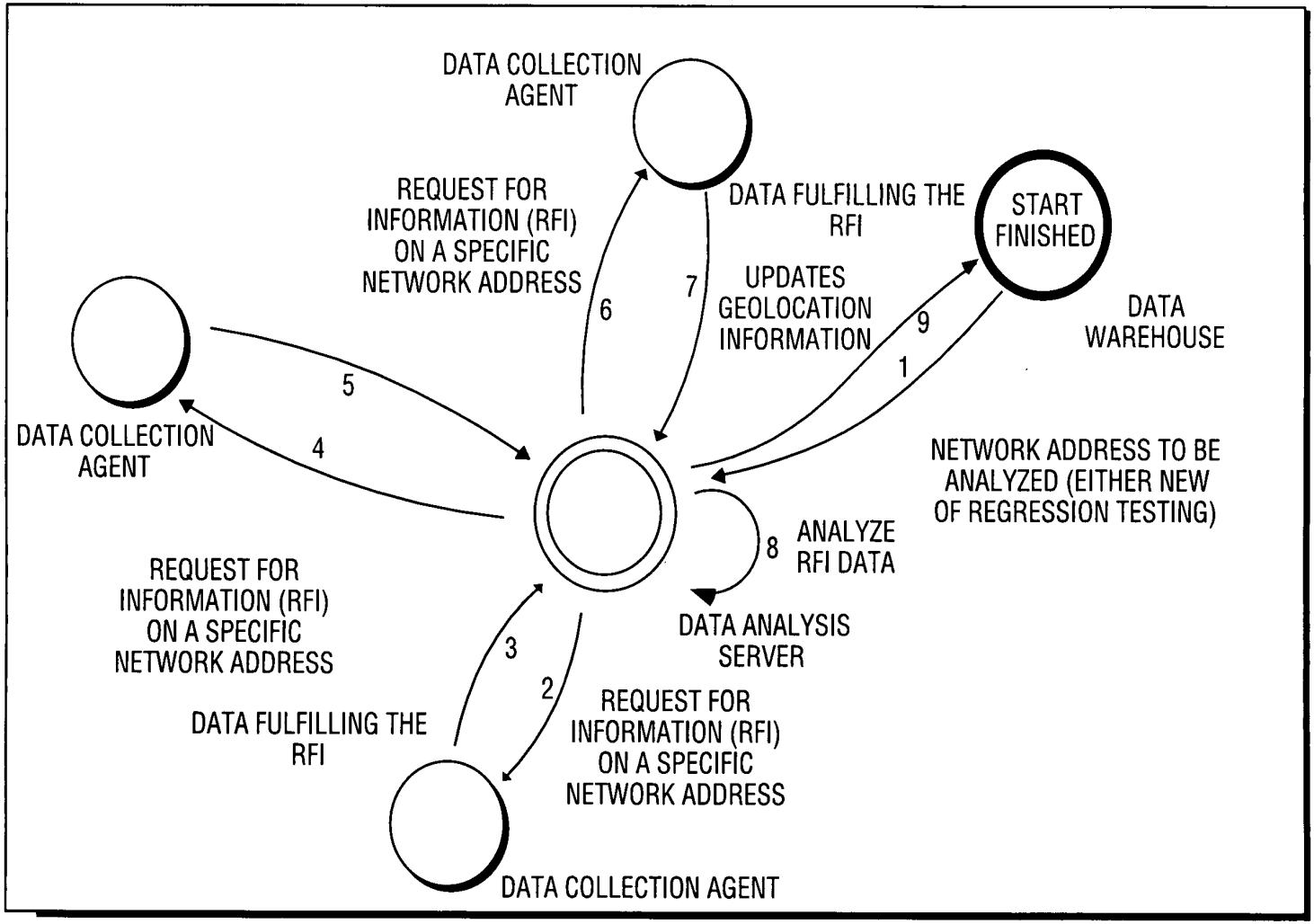

DATA COLLECTION AGENT
REQUEST FOR INFORMATION (RFI) ON A SPECIFIC NETWORK ADDRESS
DATA FULFILLING THE RFI
START FINISHED
UPDATES GEOLOCATION INFORMATION
DATA WAREHOUSE
DATA COLLECTION AGENT
NETWORK ADDRESS TO BE ANALYZED (EITHER NEW OF REGRESSION TESTING)
ANALYZE RFI DATA
REQUEST FOR INFORMATION (RFI) ON A SPECIFIC NETWORK ADDRESS
DATA ANALYSIS SERVER
DATA FULFILLING THE RFI
REQUEST FOR INFORMATION (RFI) ON A SPECIFIC NETWORK ADDRESS
DATA COLLECTION AGENT
1
2
3
4
5
6
7
8
9

FIG. 4B

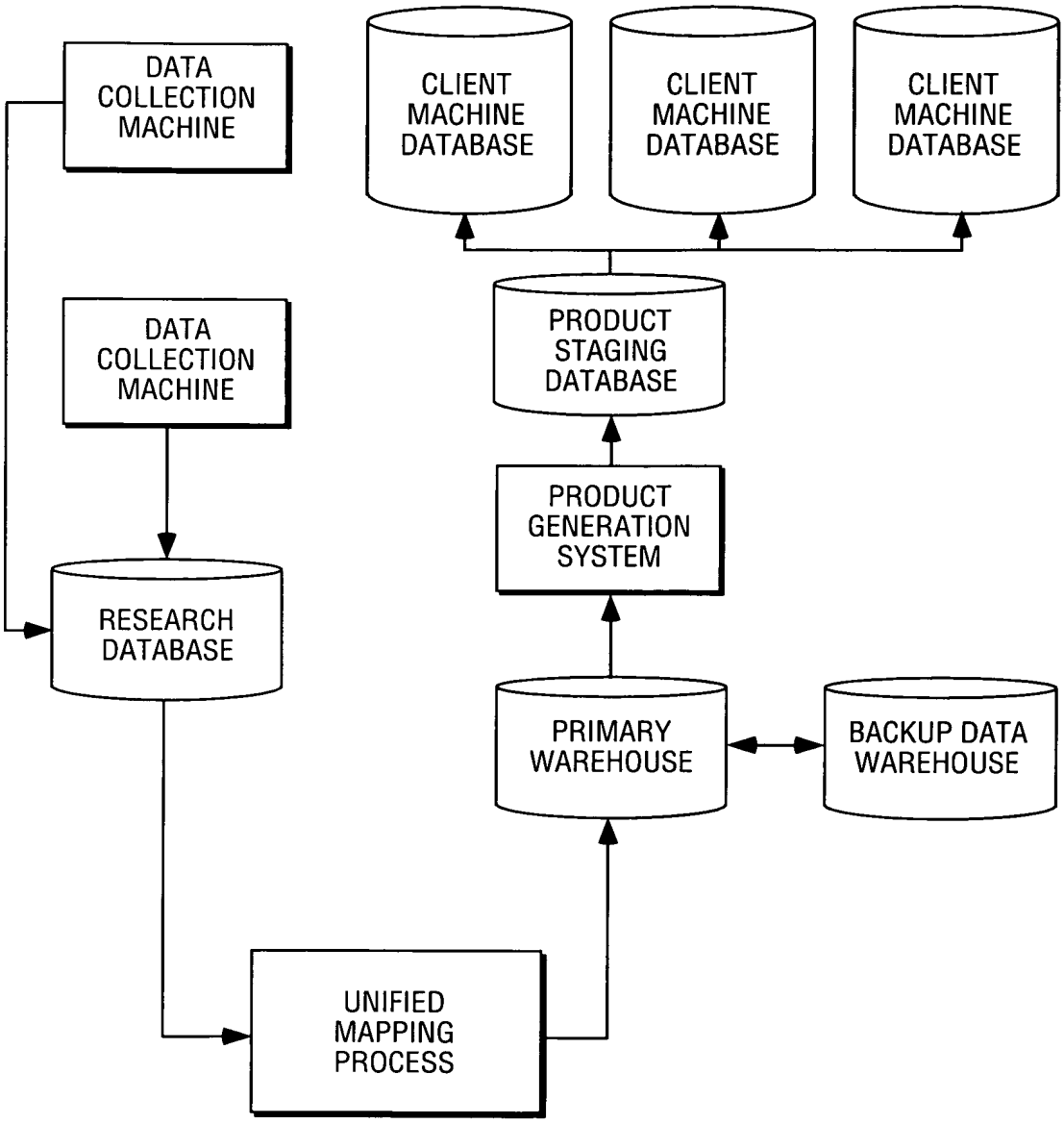

DATA COLLECTION MACHINE
DATA COLLECTION MACHINE
RESEARCH DATABASE
CLIENT MACHINE DATABASE
CLIENT MACHINE DATABASE
CLIENT MACHINE DATABASE
PRODUCT STAGING DATABASE
PRODUCT GENERATION SYSTEM
PRIMARY WAREHOUSE
BACKUP DATA WAREHOUSE
UNIFIED MAPPING PROCESS

FIG. 5

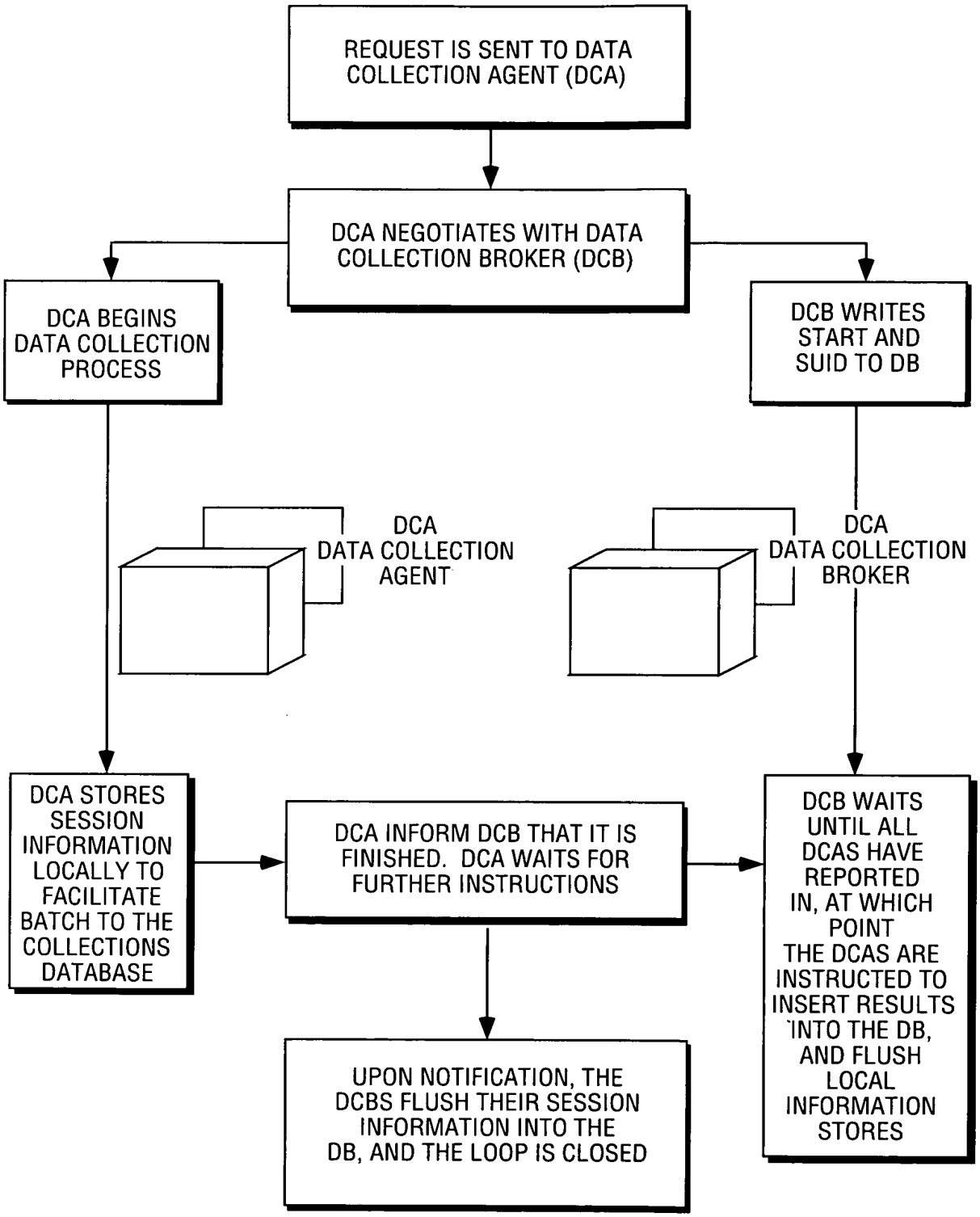

REQUEST IS SENT TO DATA COLLECTION AGENT (DCA)
DCA NEGOTIATES WITH DATA COLLECTION BROKER (DCB)
DCA BEGINS DATA COLLECTION PROCESS
DCB WRITES START AND SUID TO DB
DCA DATA COLLECTION AGENT
DCA DATA COLLECTION BROKER
DCA STORES SESSION INFORMATION LOCALLY TO FACILITATE BATCH TO THE COLLECTIONS DATABASE
DCA INFORM DCB THAT IT IS FINISHED. DCA WAITS FOR FURTHER INSTRUCTIONS
DCB WAITS UNTIL ALL DCAS HAVE REPORTED IN, AT WHICH POINT THE DCAS ARE INSTRUCTED TO INSERT RESULTS INTO THE DB, AND FLUSH LOCAL INFORMATION STORES
UPON NOTIFICATION, THE DCBS FLUSH THEIR SESSION INFORMATION INTO THE DB, AND THE LOOP IS CLOSED

FIG. 6

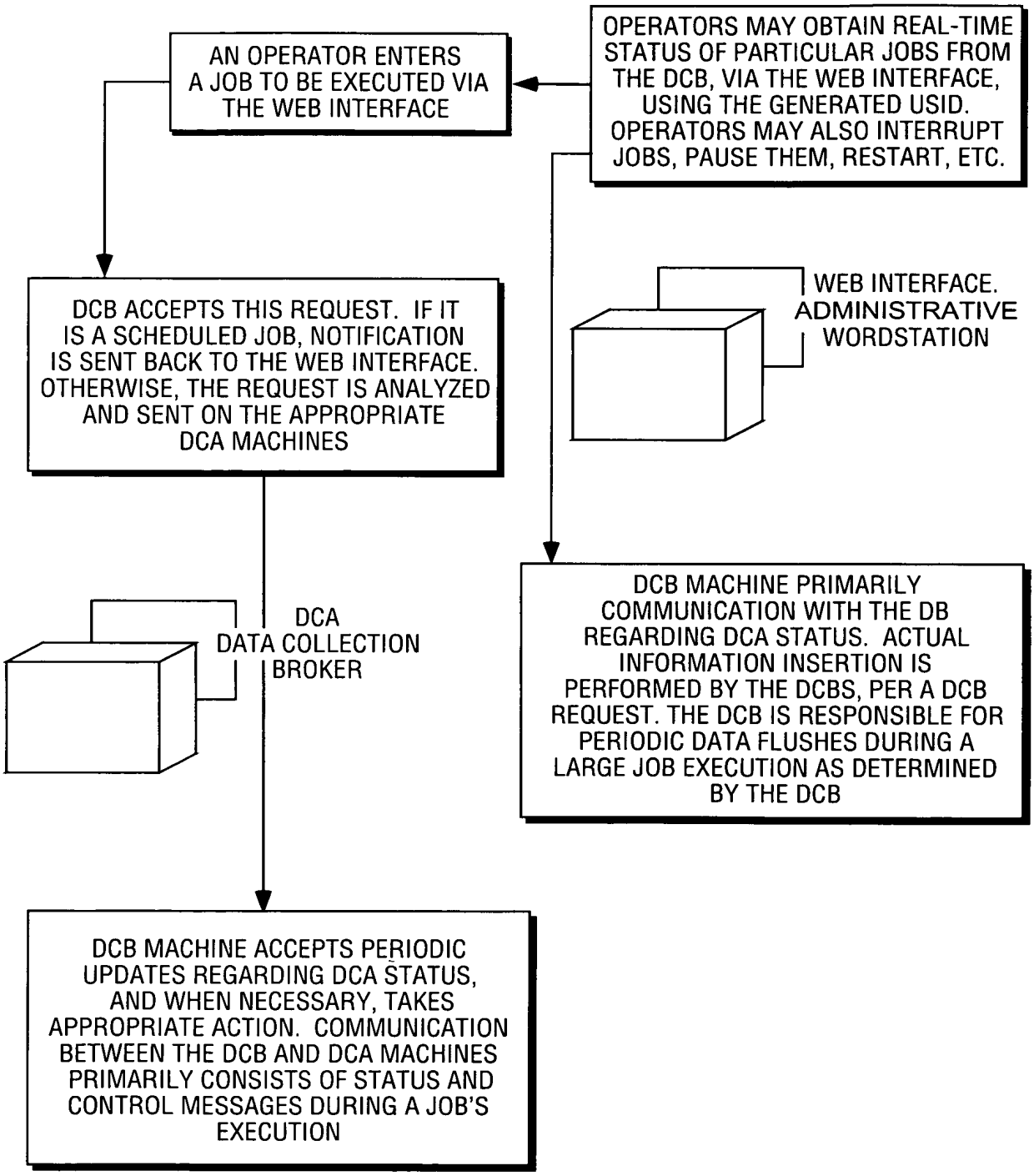

AN OPERATOR ENTERS A JOB TO BE EXECUTED VIA THE WEB INTERFACE
OPERATORS MAY OBTAIN REAL-TIME STATUS OF PARTICULAR JOBS FROM THE DCB, VIA THE WEB INTERFACE, USING THE GENERATED USID. OPERATORS MAY ALSO INTERRUPT JOBS, PAUSE THEM, RESTART, ETC.
DCB ACCEPTS THIS REQUEST. IF IT IS A SCHEDULED JOB, NOTIFICATION IS SENT BACK TO THE WEB INTERFACE. OTHERWISE, THE REQUEST IS ANALYZED AND SENT ON THE APPROPRIATE DCA MACHINES
WEB INTERFACE. ADMINISTRATIVE WORDSTATION
DCA DATA COLLECTION BROKER
DCB MACHINE PRIMARILY COMMUNICATION WITH THE DB REGARDING DCA STATUS. ACTUAL INFORMATION INSERTION IS PERFORMED BY THE DCBS, PER A DCB REQUEST. THE DCB IS RESPONSIBLE FOR PERIODIC DATA FLUSHES DURING A LARGE JOB EXECUTION AS DETERMINED BY THE DCB
DCB MACHINE ACCEPTS PERIODIC UPDATES REGARDING DCA STATUS, AND WHEN NECESSARY, TAKES APPROPRIATE ACTION. COMMUNICATION BETWEEN THE DCB AND DCA MACHINES PRIMARILY CONSISTS OF STATUS AND CONTROL MESSAGES DURING A JOB'S EXECUTION

FIG. 7

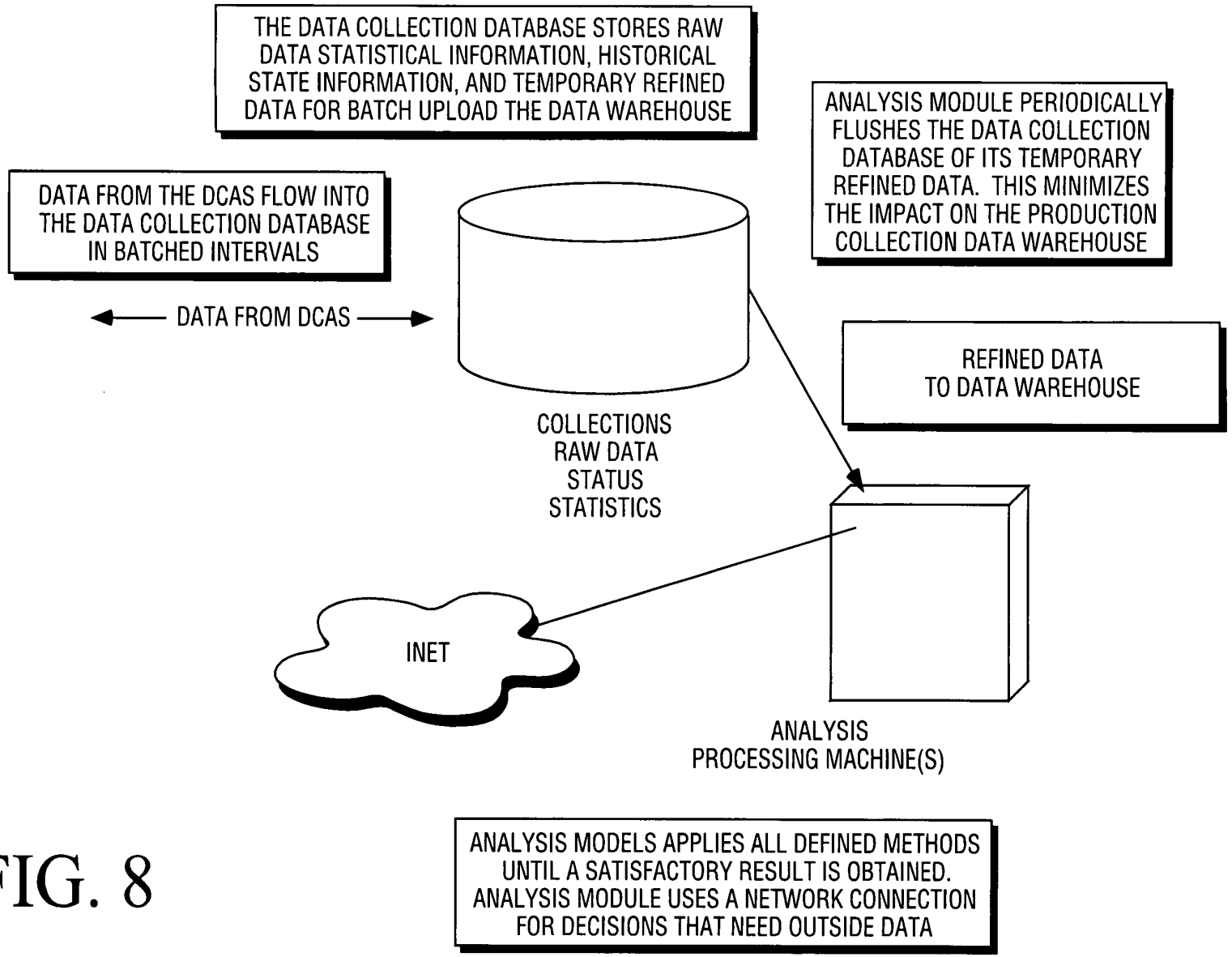

THE DATA COLLECTION DATABASE STORES RAW DATA STATISTICAL INFORMATION, HISTORICAL STATE INFORMATION, AND TEMPORARY REFINED DATA FOR BATCH UPLOAD THE DATA WAREHOUSE
ANALYSIS MODULE PERIODICALLY FLUSHES THE DATA COLLECTION DATABASE OF ITS TEMPORARY REFINED DATA. THIS MINIMIZES THE IMPACT ON THE PRODUCTION COLLECTION DATA WAREHOUSE
DATA FROM THE DCAS FLOW INTO THE DATA COLLECTION DATABASE IN BATCHED INTERVALS
DATA FROM DCAS
REFINED DATA TO DATA WAREHOUSE
COLLECTIONS RAW DATA STATUS STATISTICS
INET
ANALYSIS PROCESSING MACHINE(S)
ANALYSIS MODELS APPLIES ALL DEFINED METHODS UNTIL A SATISFACTORY RESULT IS OBTAINED. ANALYSIS MODULE USES A NETWORK CONNECTION FOR DECISIONS THAT NEED OUTSIDE DATA

FIG. 8

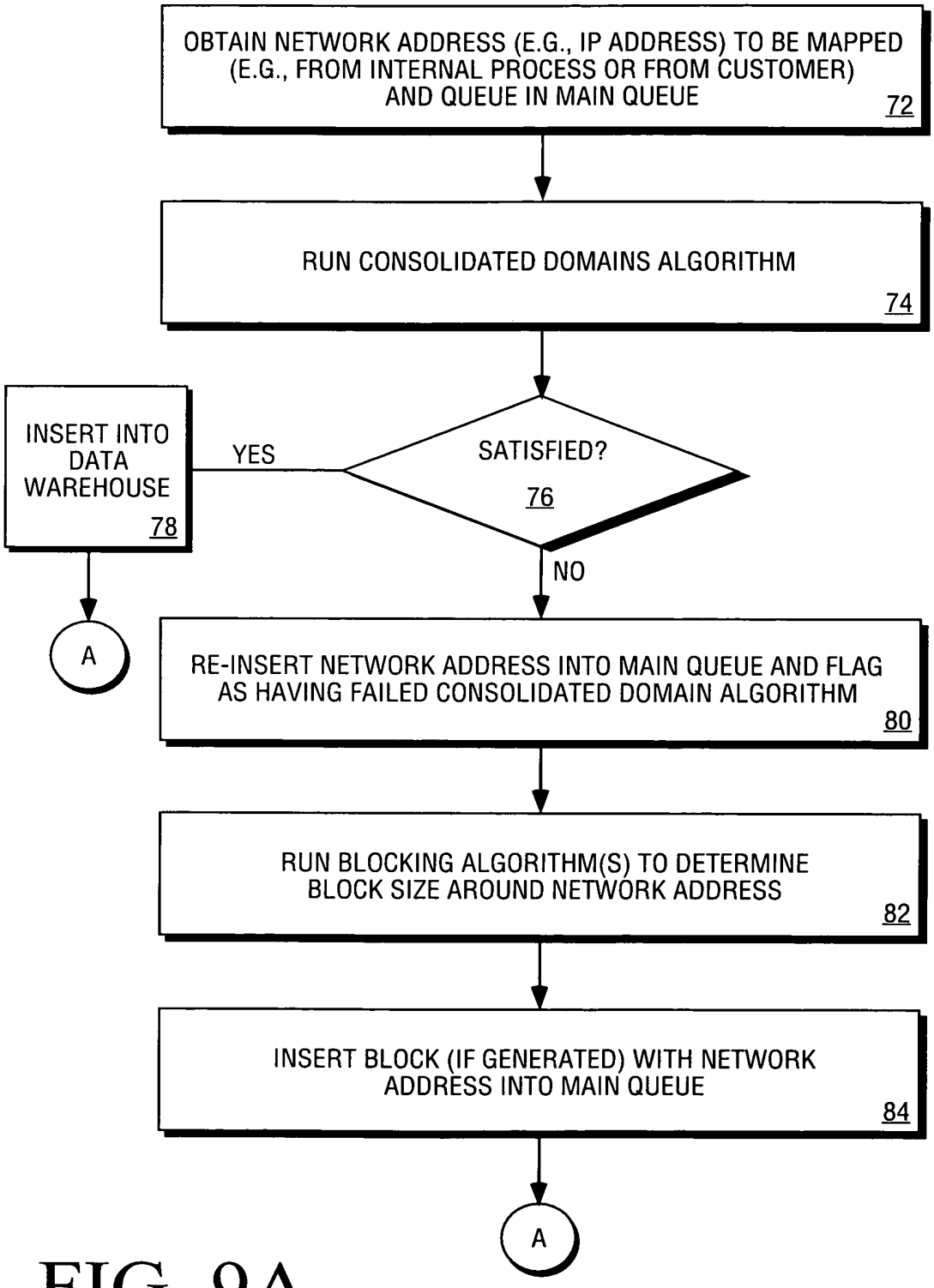

OBTAIN NETWORK ADDRESS (E.G., IP ADDRESS) TO BE MAPPED (E.G., FROM INTERNAL PROCESS OR FROM CUSTOMER) AND QUEUE IN MAIN QUEUE
72
RUN CONSOLIDATED DOMAINS ALGORITHM
74
SATISFIED?
76
YES
INSERT INTO DATA WAREHOUSE
78
A
NO
RE-INSERT NETWORK ADDRESS INTO MAIN QUEUE AND FLAG AS HAVING FAILED CONSOLIDATED DOMAIN ALGORITHM
80
RUN BLOCKING ALGORITHM(S) TO DETERMINE BLOCK SIZE AROUND NETWORK ADDRESS
82
INSERT BLOCK (IF GENERATED) WITH NETWORK ADDRESS INTO MAIN QUEUE
84
A

FIG. 9A

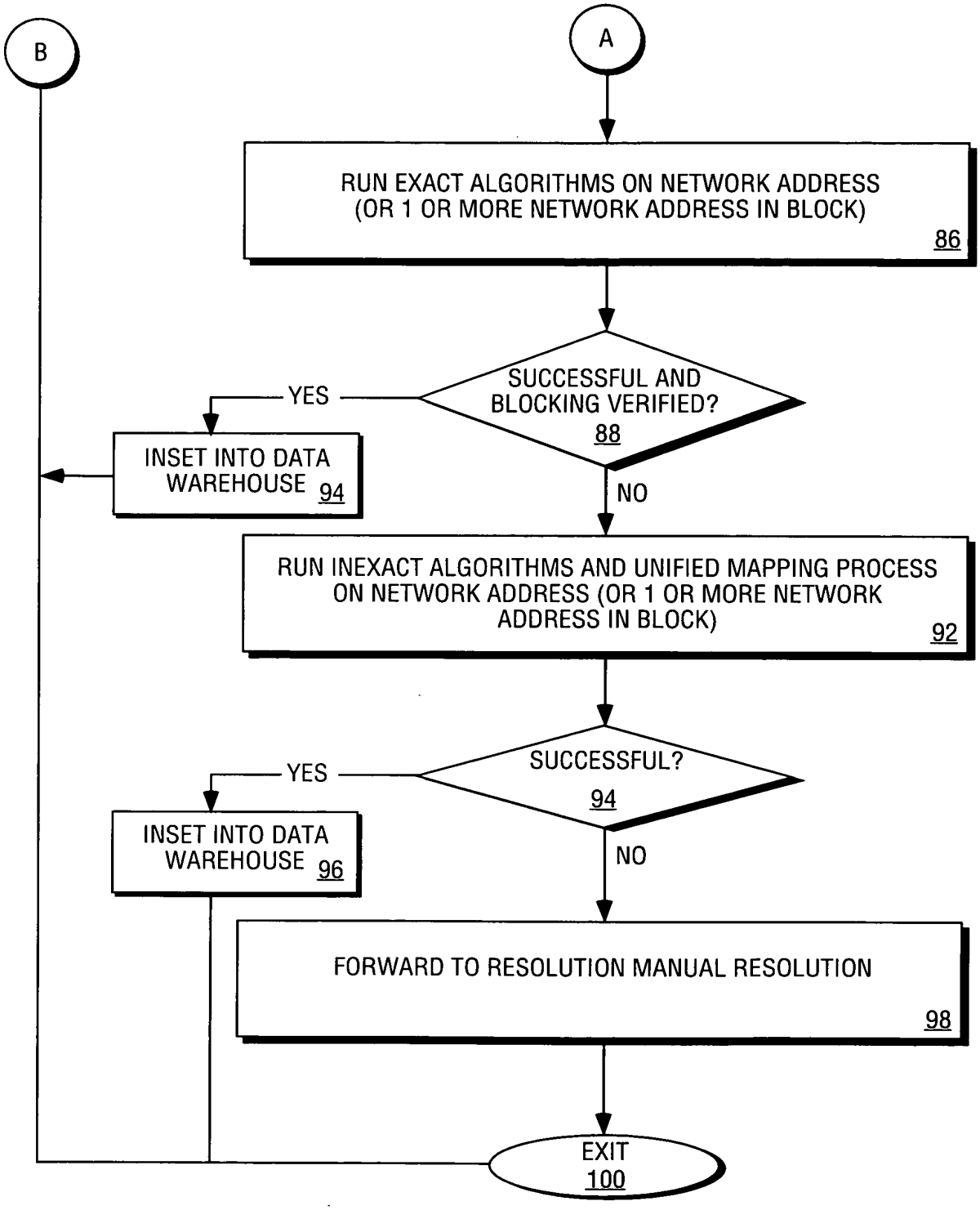

B
A
RUN EXACT ALGORITHMS ON NETWORK ADDRESS (OR 1 OR MORE NETWORK ADDRESS IN BLOCK)
86
SUCCESSFUL AND BLOCKING VERIFIED?
88
YES
INSET INTO DATA WAREHOUSE 94
NO
RUN INEXACT ALGORITHMS AND UNIFIED MAPPING PROCESS ON NETWORK ADDRESS (OR 1 OR MORE NETWORK ADDRESS IN BLOCK)
92
SUCCESSFUL?
94
YES
INSET INTO DATA WAREHOUSE 96
NO
FORWARD TO RESOLUTION MANUAL RESOLUTION
98
EXIT
100

FIG. 9B

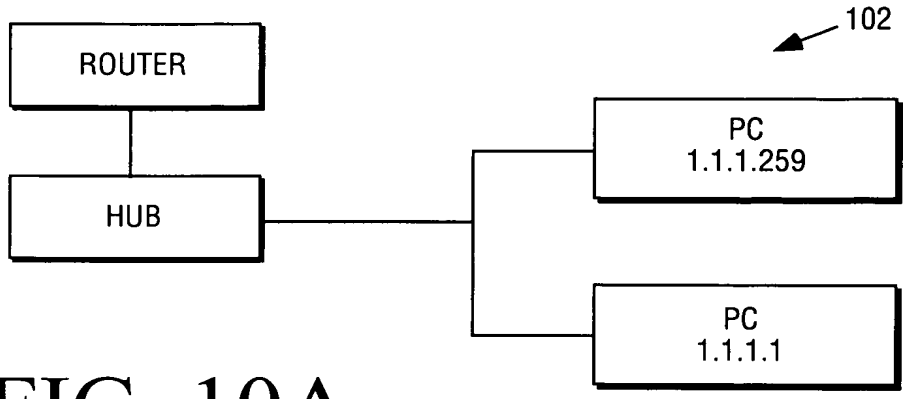


FIG. 10A

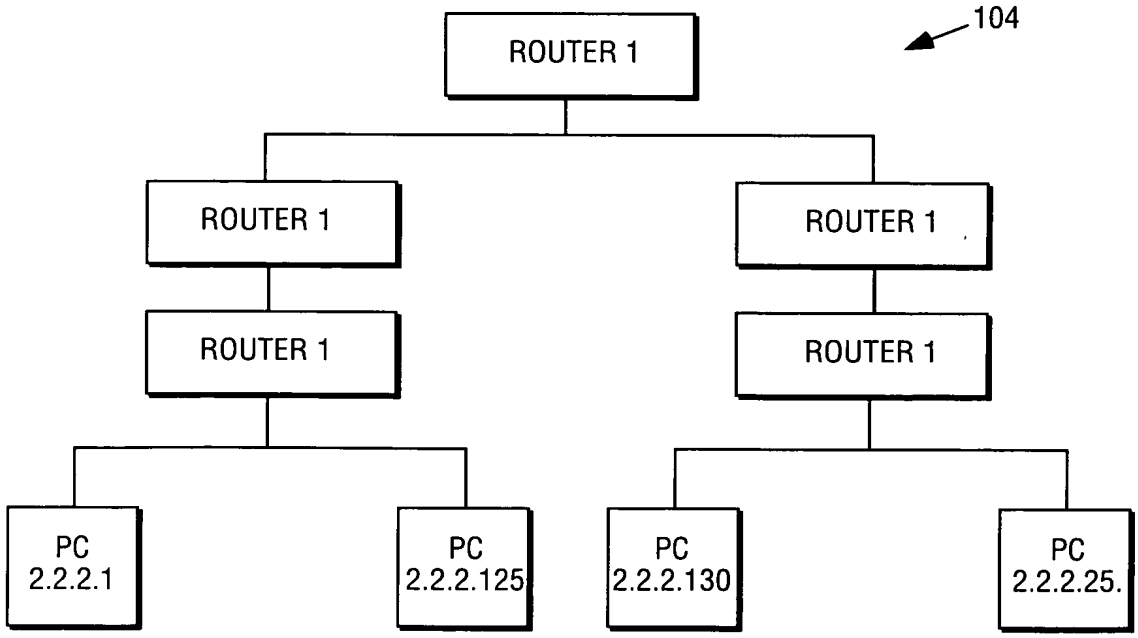


FIG. 10B

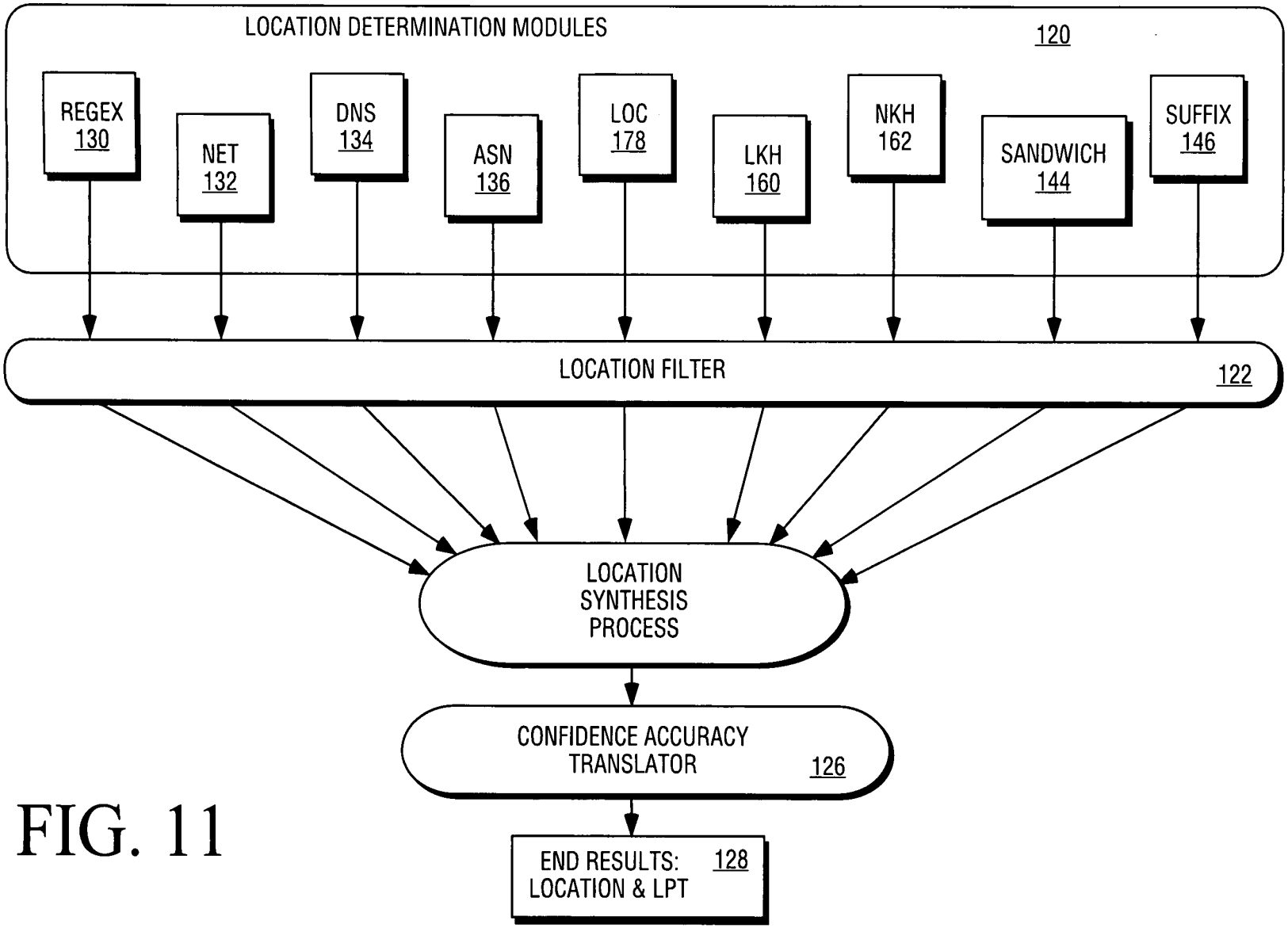

LOCATION DETERMINATION MODULES
120
REGEX
130
NET
132
DNS
134
ASN
136
LOC
178
LKH
160
NKH
162
SANDWICH
144
SUFFIX
146
LOCATION FILTER
122
LOCATION
SYNTHESIS
PROCESS
CONFIDENCE ACCURACY
TRANSLATOR
126
END RESULTS:
LOCATION & LPT
128

FIG. 11

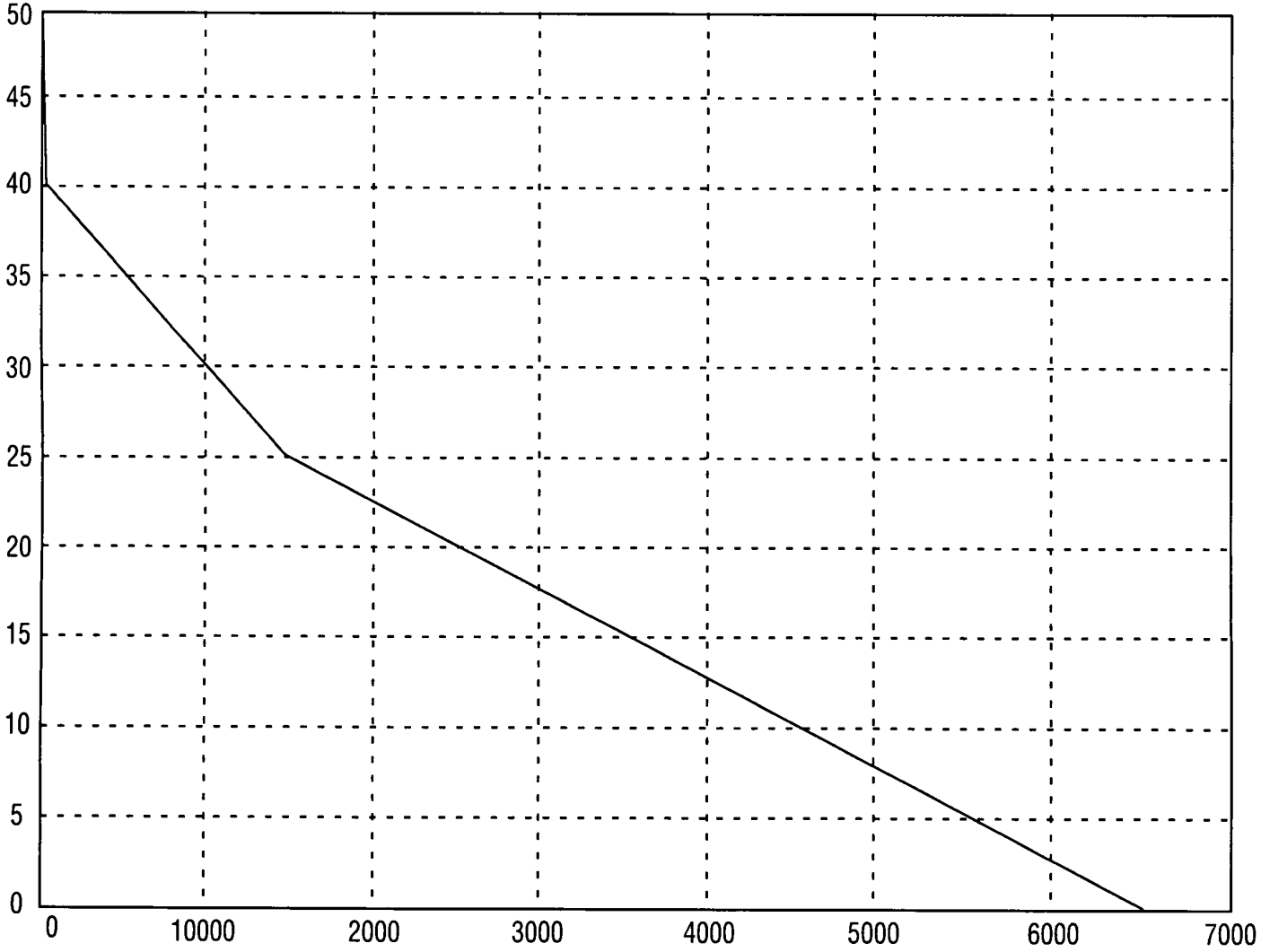

150
50
45
40
35
30
25
20
15
10
5
0
0
10000
2000
3000
4000
5000
6000
7000
NUMBER OF NODES IN NETWORK BLOCK

FIG. 12A

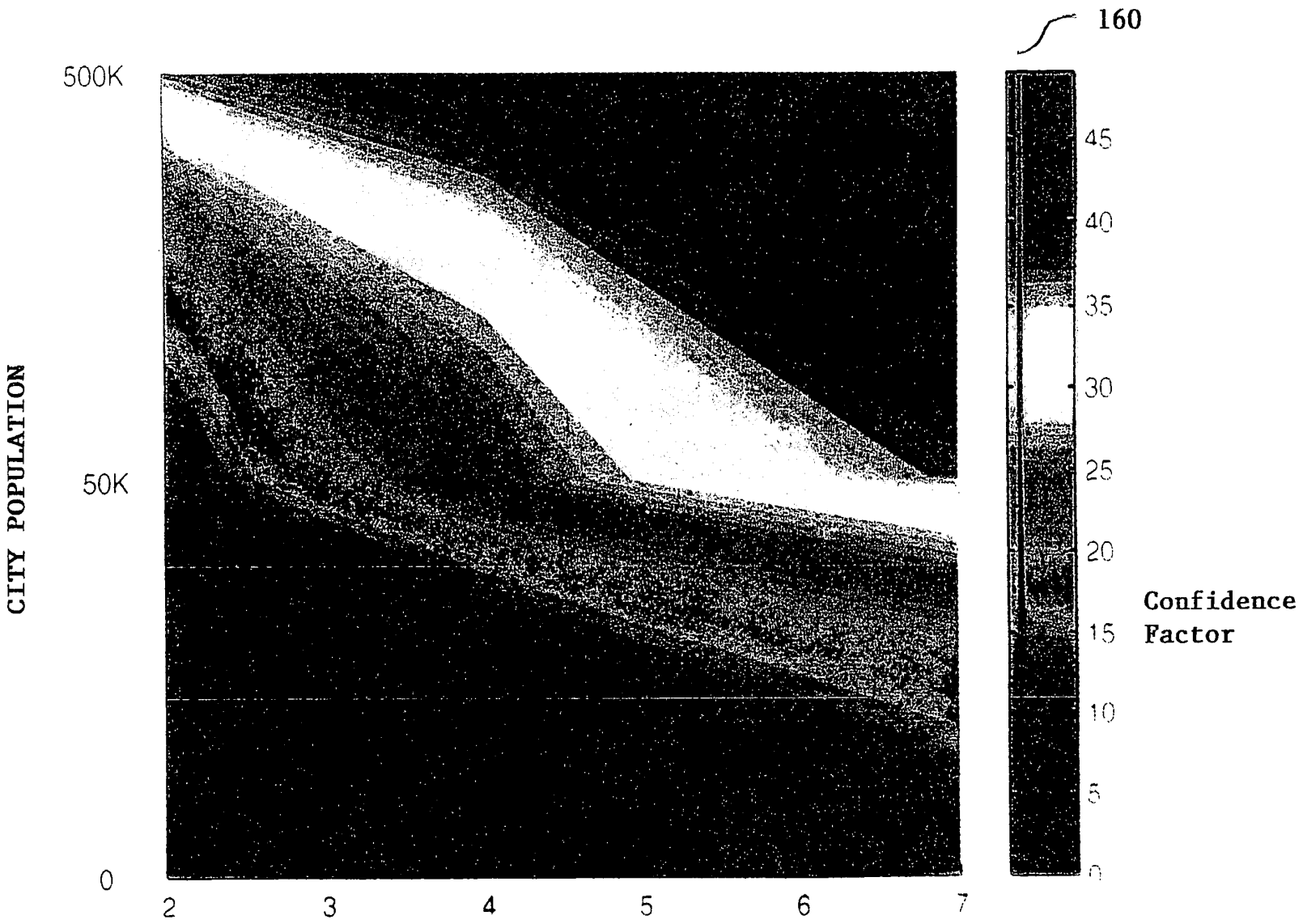

160
500K
50K
0
CITY POPULATION
2
3
4
5
6
7
String Length
45
40
35
30
25
20
15
10
5
0
Confidence
Factor

FIG. 12B

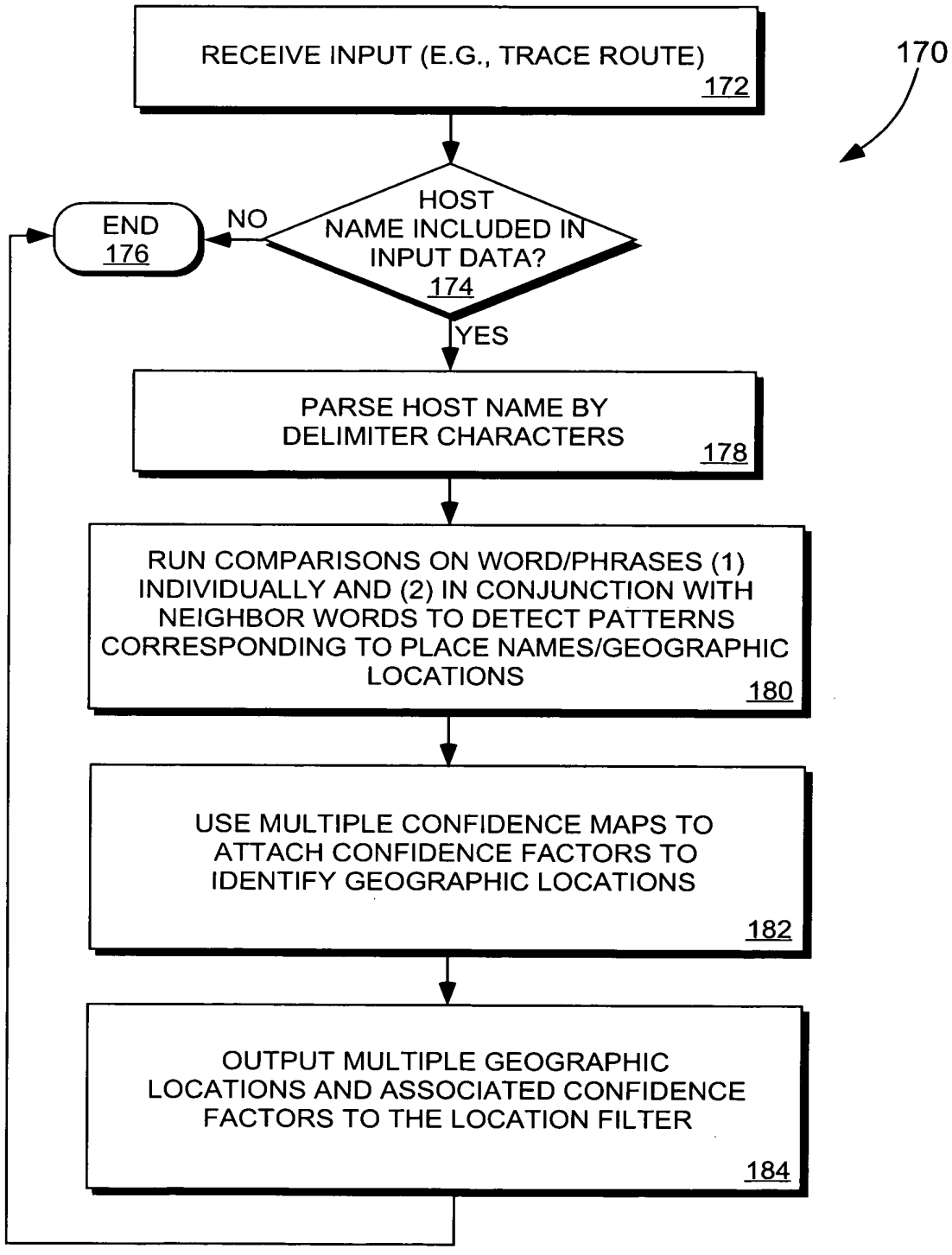

RECEIVE INPUT (E.G., TRACE ROUTE)
172
170
HOST NAME INCLUDED IN INPUT DATA?
174
NO
END
176
YES
PARSE HOST NAME BY DELIMITER CHARACTERS
178
RUN COMPARISONS ON WORD/PHRASES (1) INDIVIDUALLY AND (2) IN CONJUNCTION WITH NEIGHBOR WORDS TO DETECT PATTERNS CORRESPONDING TO PLACE NAMES/GEOGRAPHIC LOCATIONS
180
USE MULTIPLE CONFIDENCE MAPS TO ATTACH CONFIDENCE FACTORS TO IDENTIFY GEOGRAPHIC LOCATIONS
182
OUTPUT MULTIPLE GEOGRAPHIC LOCATIONS AND ASSOCIATED CONFIDENCE FACTORS TO THE LOCATION FILTER
184

FIG. 13

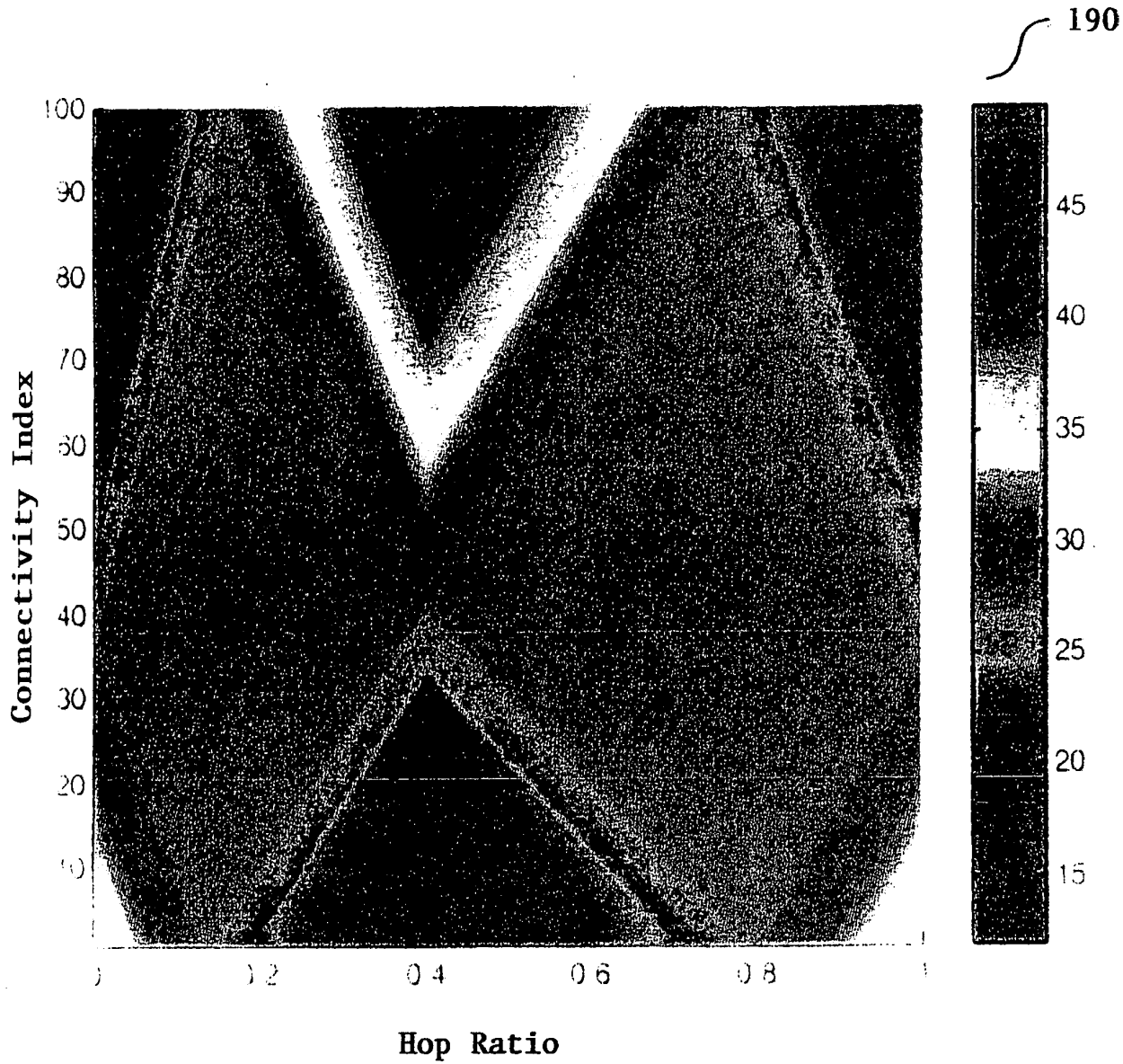

190
100
90
80
70
60
50
40
30
20
10
Connectivity Index
0
0.2
0.4
0.6
0.8
1
Hop Ratio
45
40
35
30
25
20
15

FIG. 14A

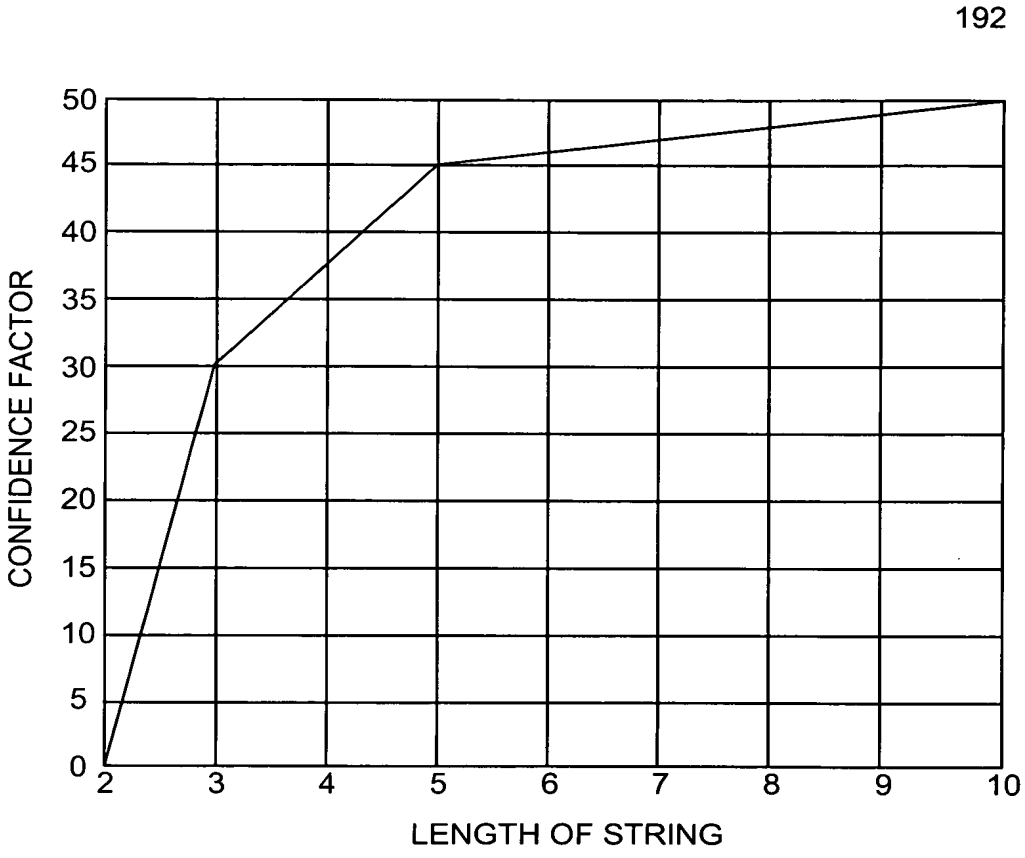

192
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
2
3
4
5
6
7
8
9
10
LENGTH OF STRING

FIG. 14B

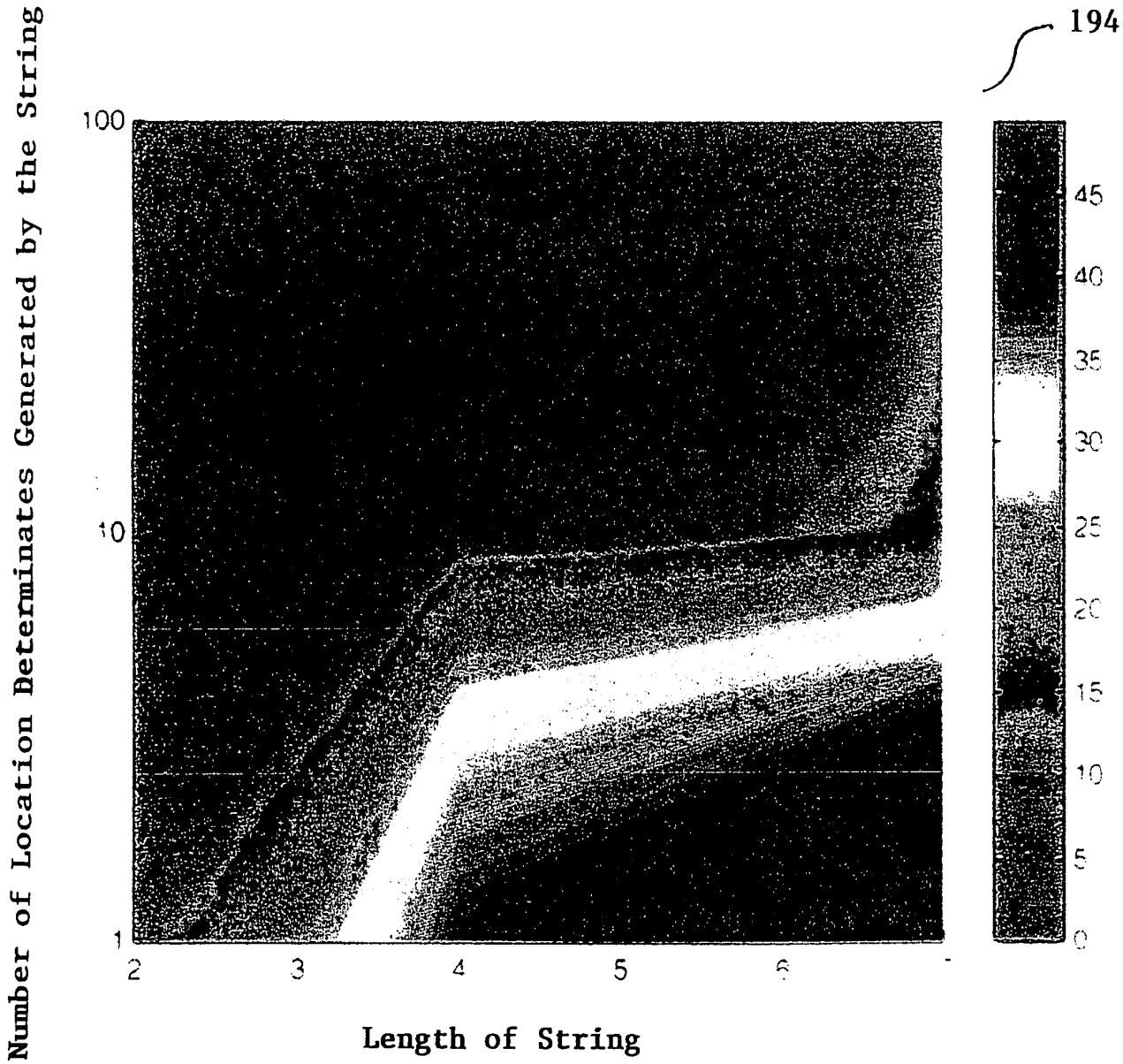

194
Number of Location Determinates Generated by the String
100
10
1
2
3
4
5
6
Length of String
45
40
35
30
25
20
15
10
5
0

FIG. 14C

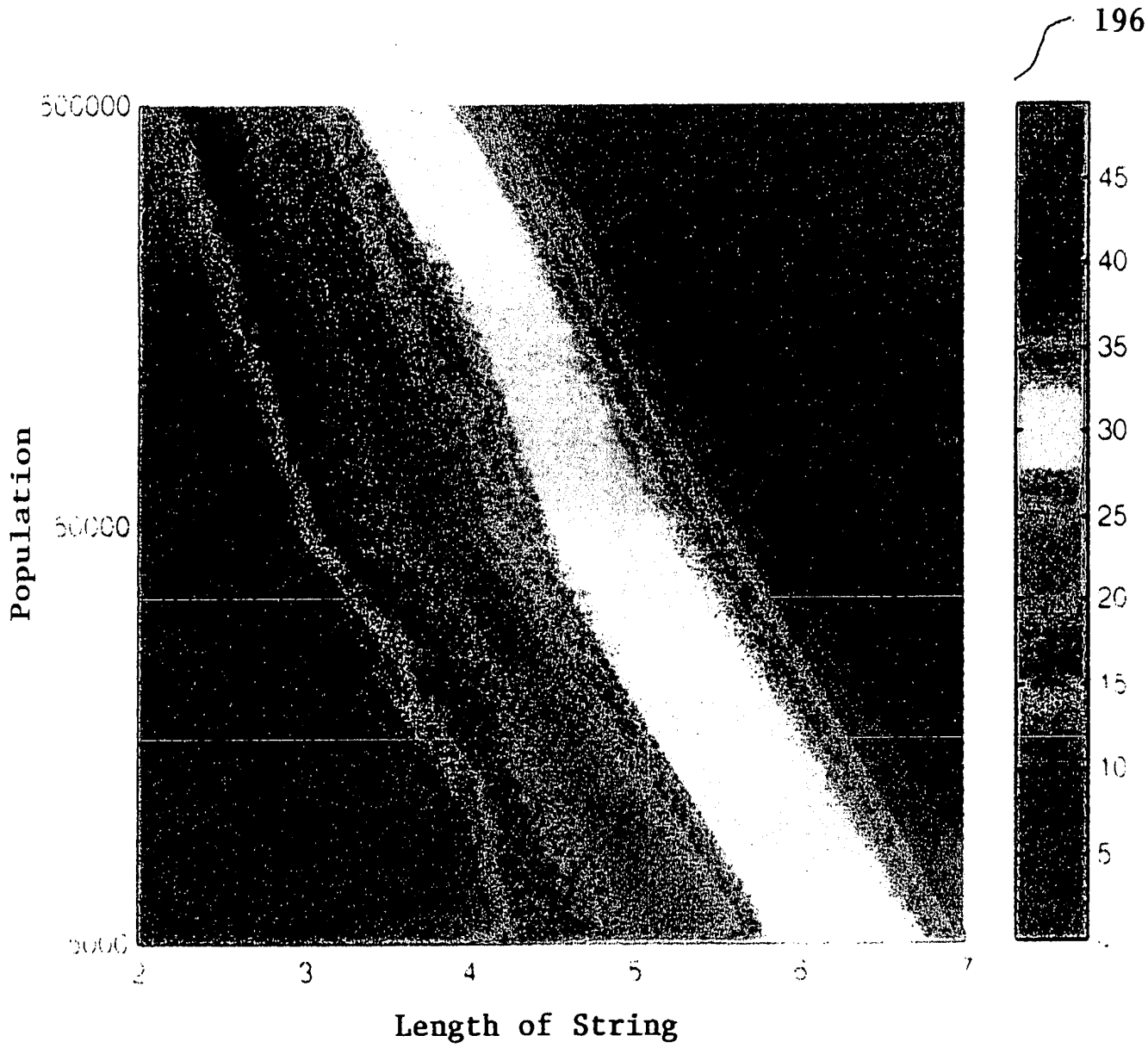

196
500000
50000
2000
Population
2
3
4
5
6
7
Length of String
45
40
35
30
25
20
15
10
5

FIG. 14D

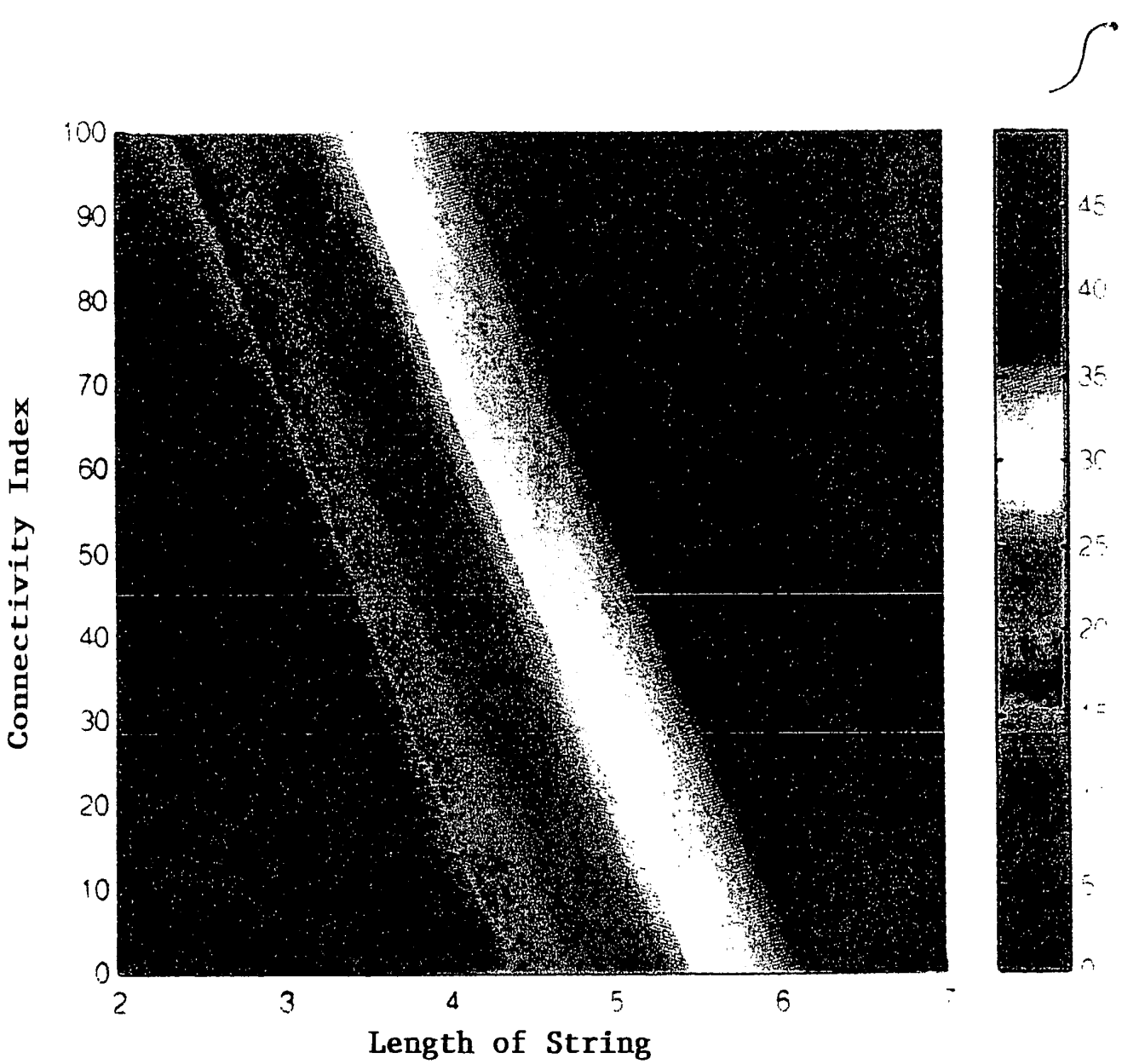

198
Connectivity Index
100
90
80
70
60
50
40
30
20
10
0
2
3
4
5
6
Length of String
45
40
35
30

FIG. 14E

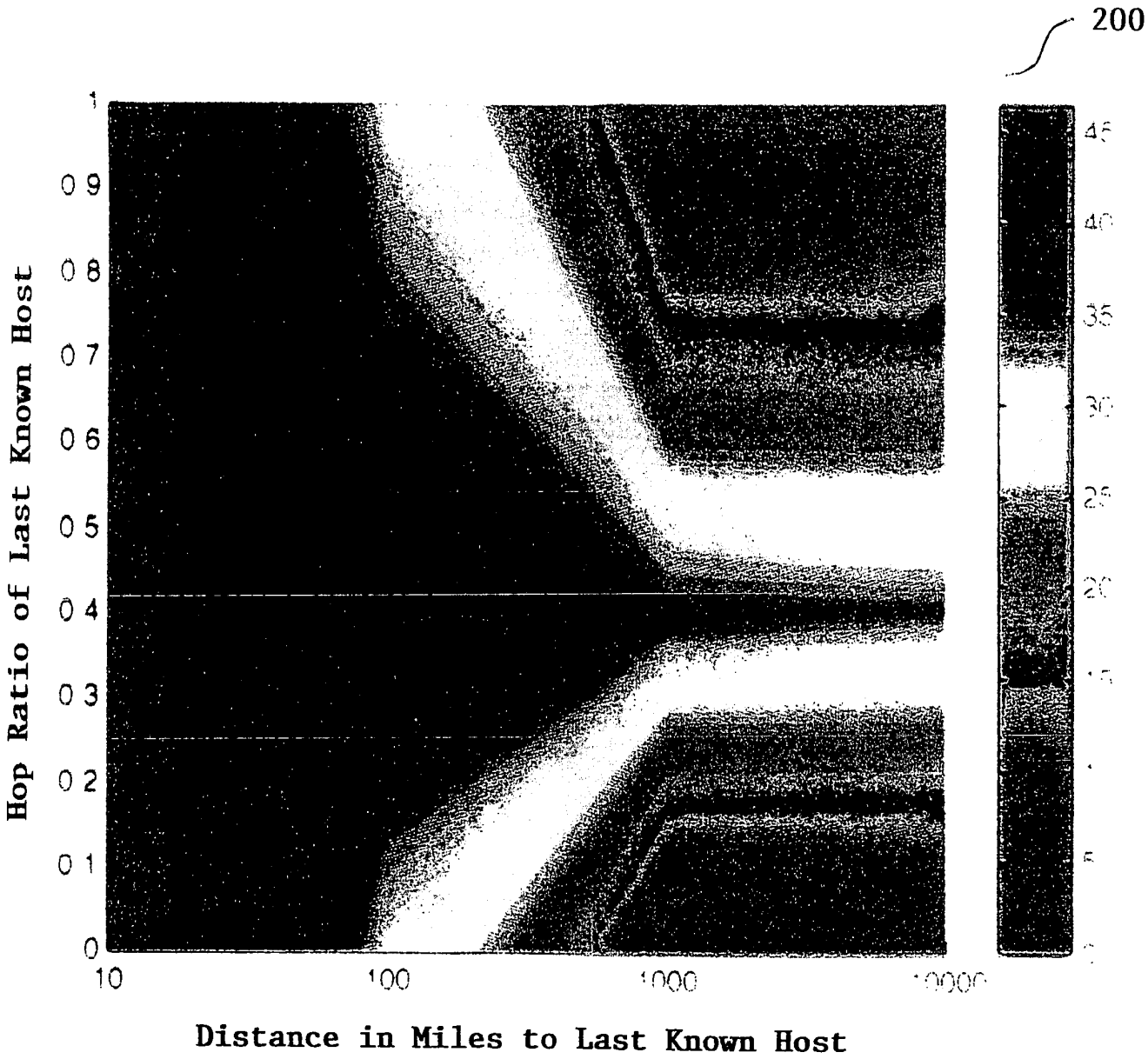

200
Hop Ratio of Last Known Host
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
10
100
1000
Distance in Miles to Last Known Host

FIG. 14F

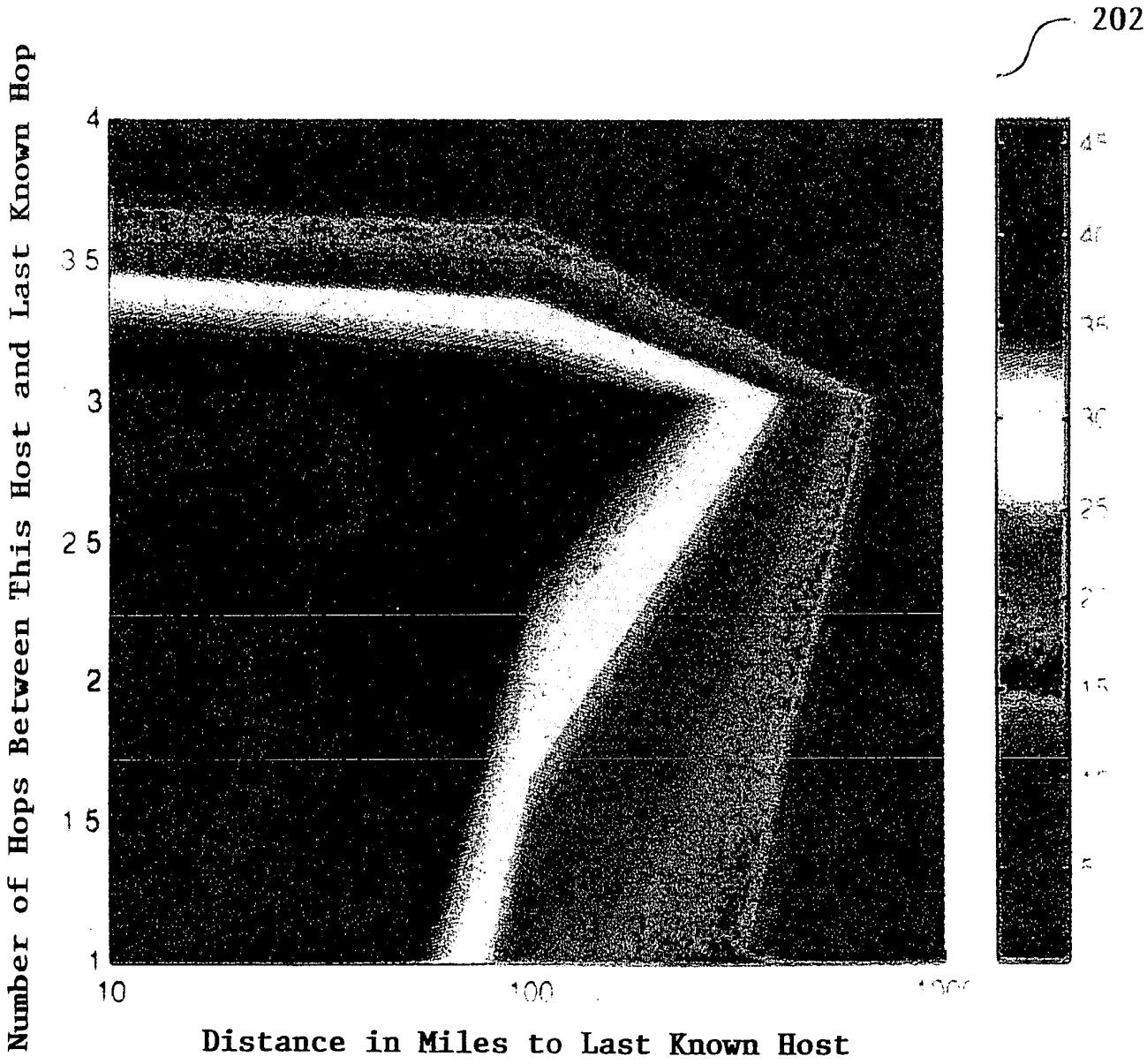

202
Number of Hops Between This Host and Last Known Hop
4
3.5
3
2.5
2
1.5
1
10
100
Distance in Miles to Last Known Host

FIG. 14G

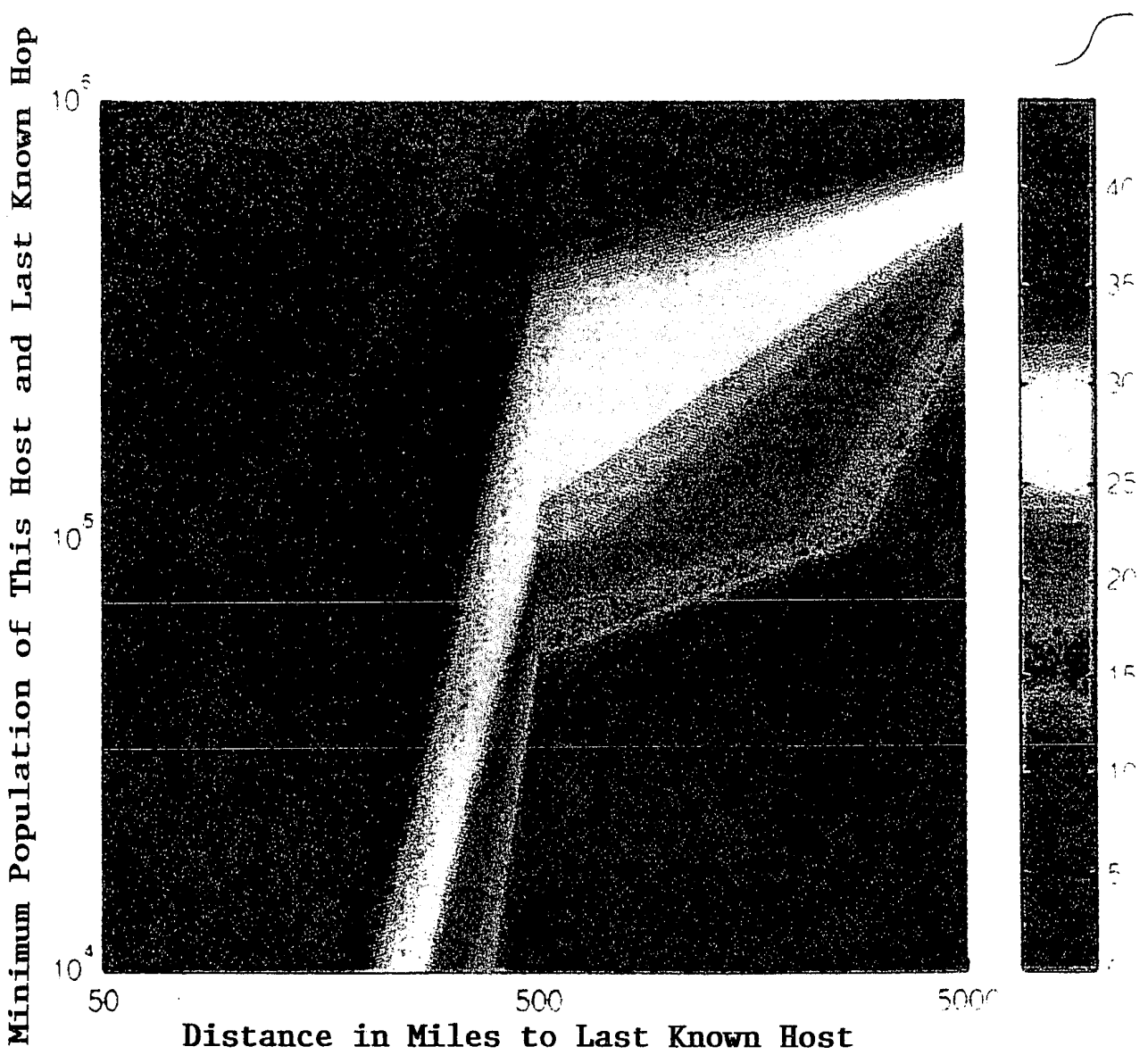

204
Minimum Population of This Host and Last Known Hop
10^6
10^5
10^4
50
500
5000
Distance in Miles to Last Known Host

FIG. 14H

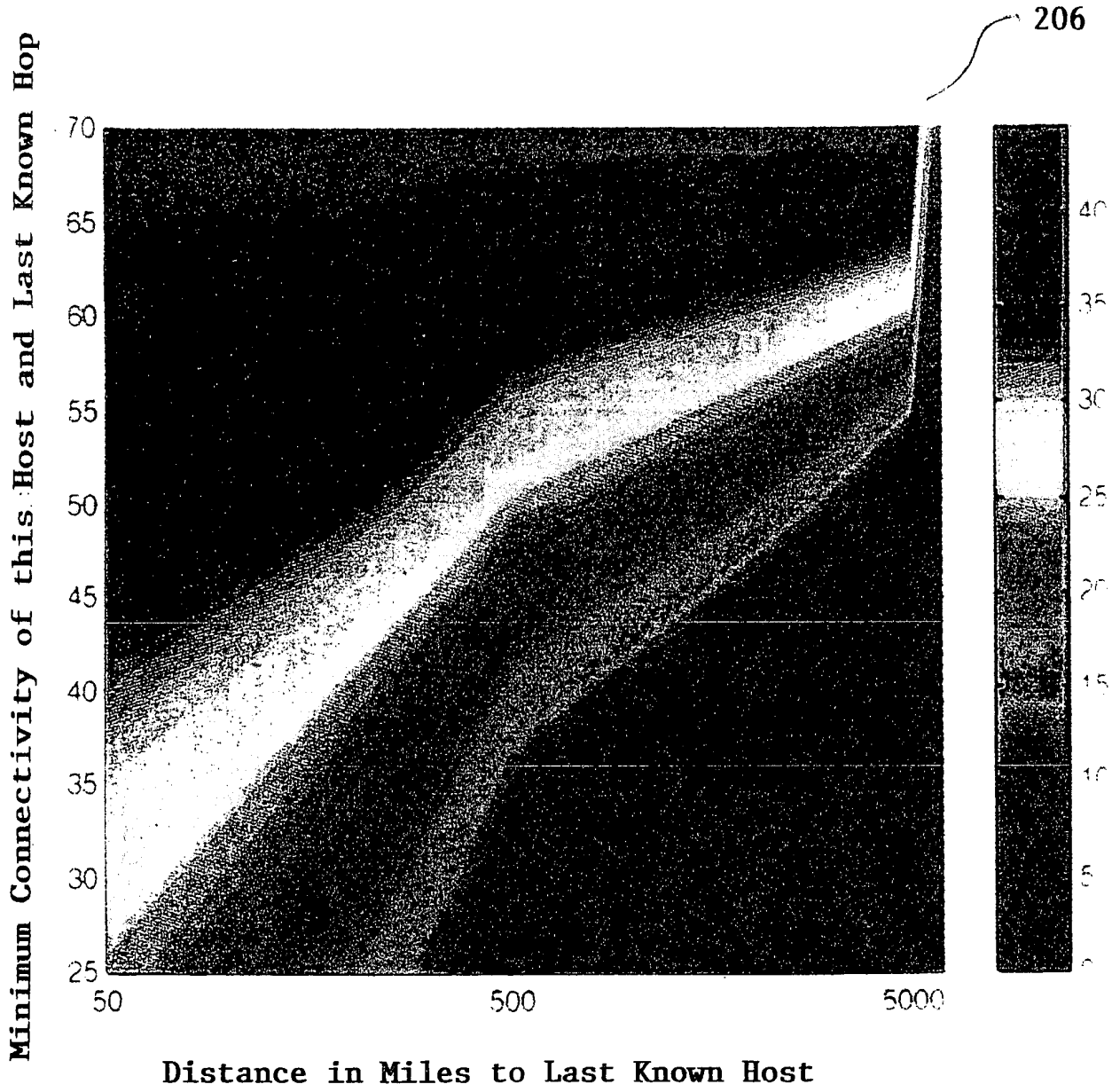

206
Minimum Connectivity of this Host and Last Known Hop
70
65
60
55
50
45
40
35
30
25
50
500
5000
40
35
30
25
20
15
10
5
0
Distance in Miles to Last Known Host

FIG. 14I

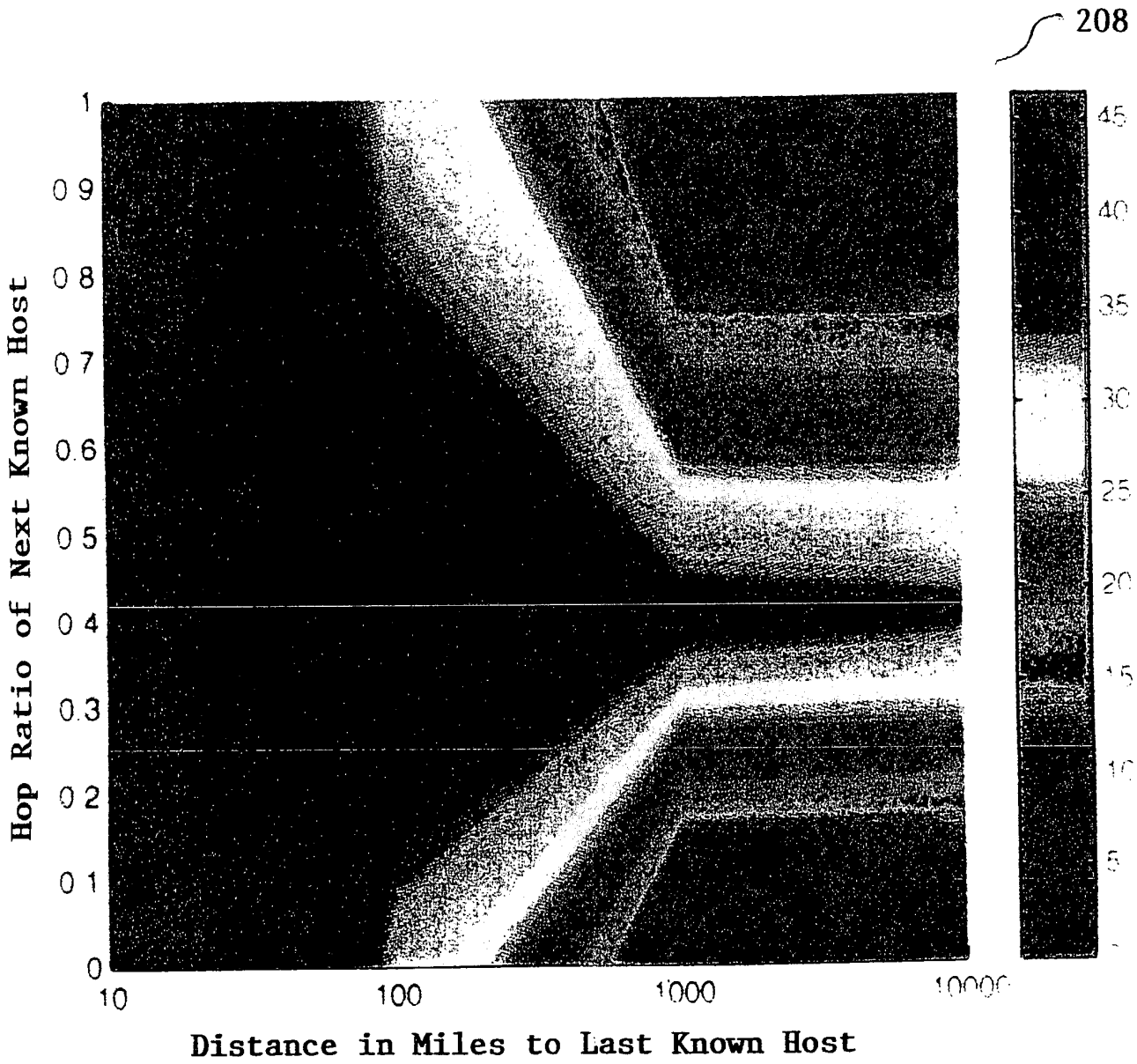

208
Hop Ratio of Next Known Host
1
0 9
0 8
0 7
0.6
0 5
0 4
0.3
0 2
0 1
0
10
100
1000
10000
Distance in Miles to Last Known Host
45
40
35
30
25
20
15
10
5

FIG. 14J

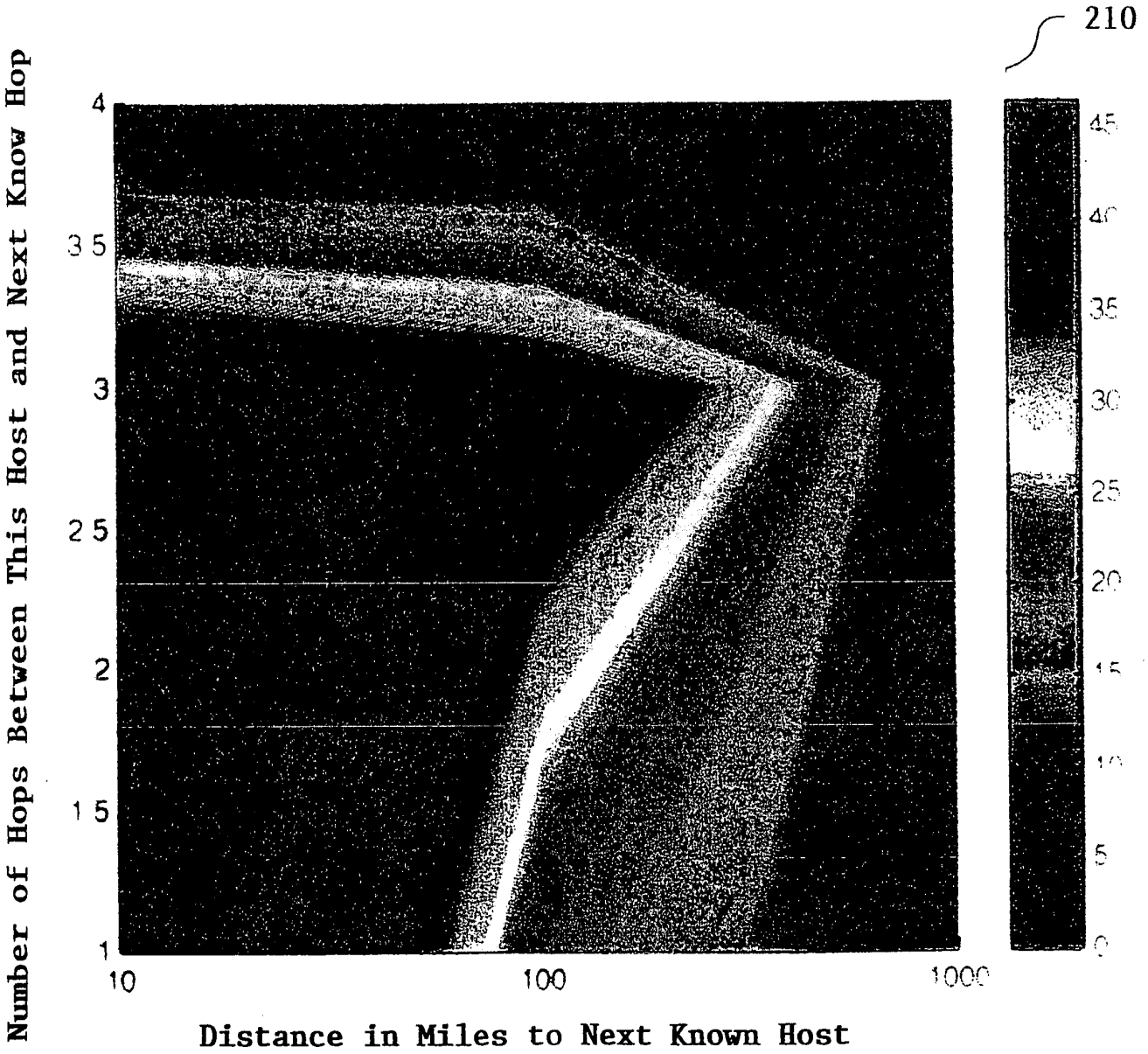

210
Number of Hops Between This Host and Next Know Hop
4
3 5
3
2 5
2
1 5
1
10
100
1000
45
40
35
30
25
20
15
10
5
0
Distance in Miles to Next Known Host

FIG. 14K

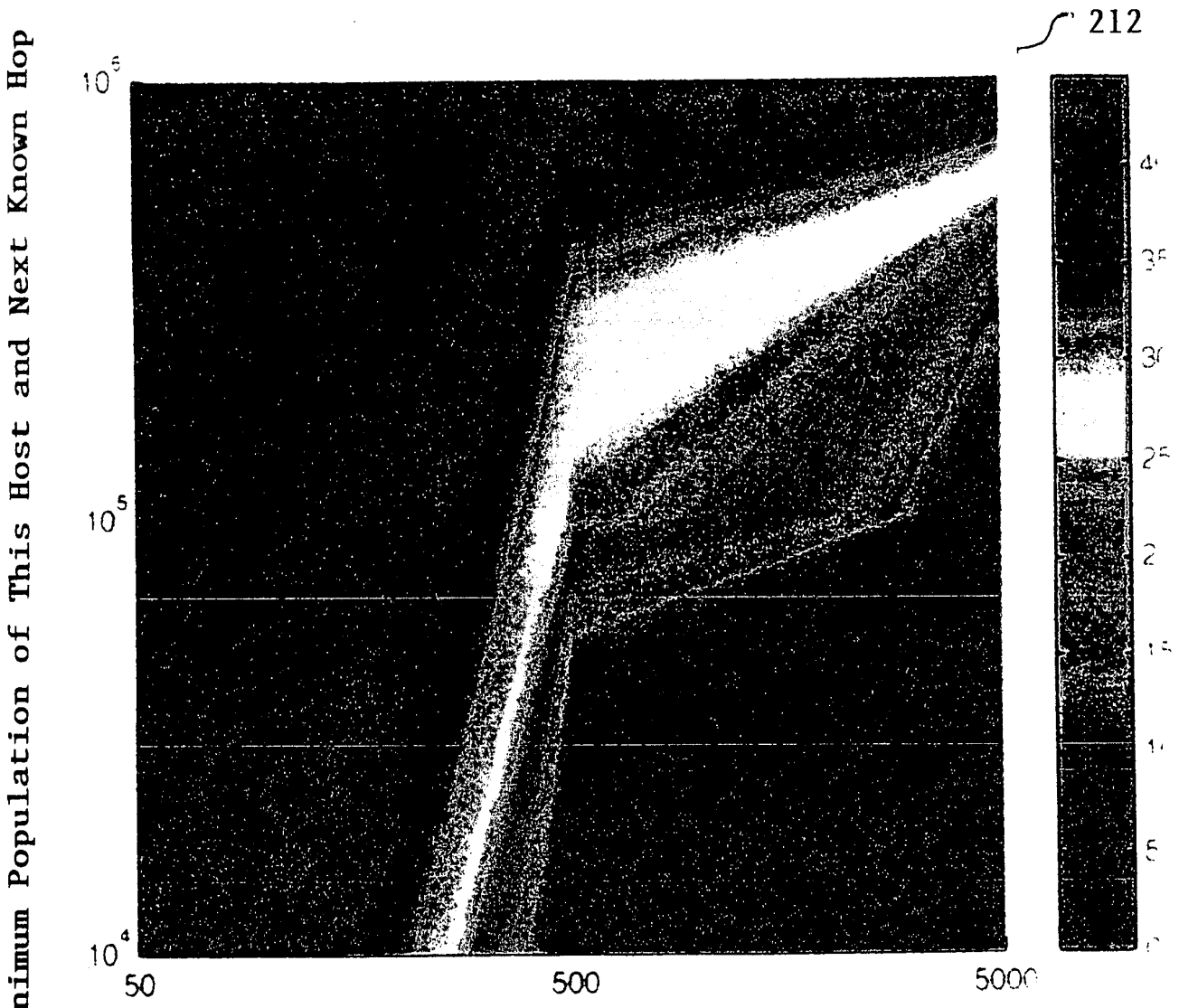

212
Minimum Population of This Host and Next Known Hop
10^6
10^5
10^4
50
500
5000
Distance in Miles to Next Known Host

FIG. 14L

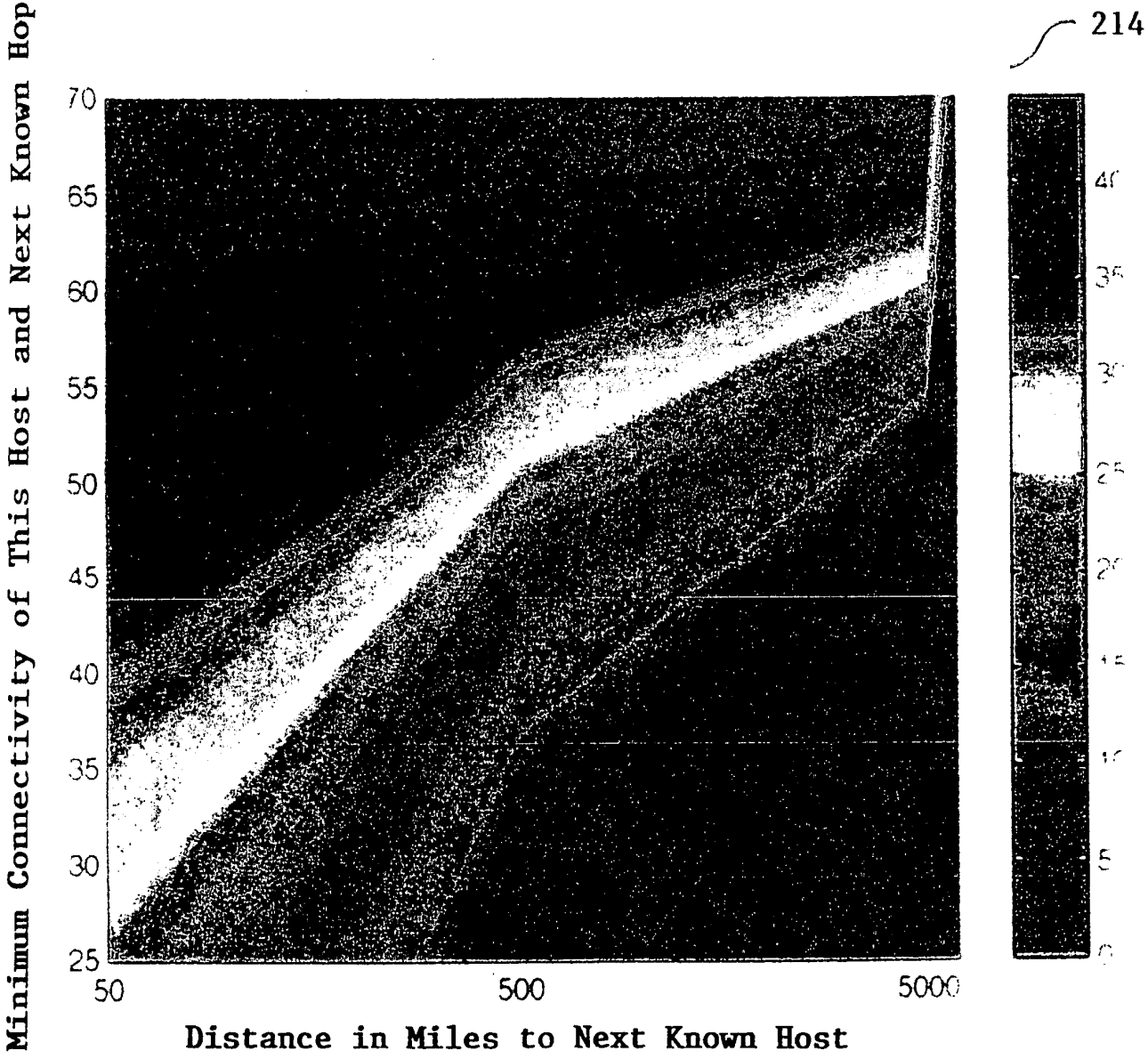

214
Minimum Connectivity of This Host and Next Known Hop
70
65
60
55
50
45
40
35
30
25
50
500
5000
Distance in Miles to Next Known Host

FIG. 14M

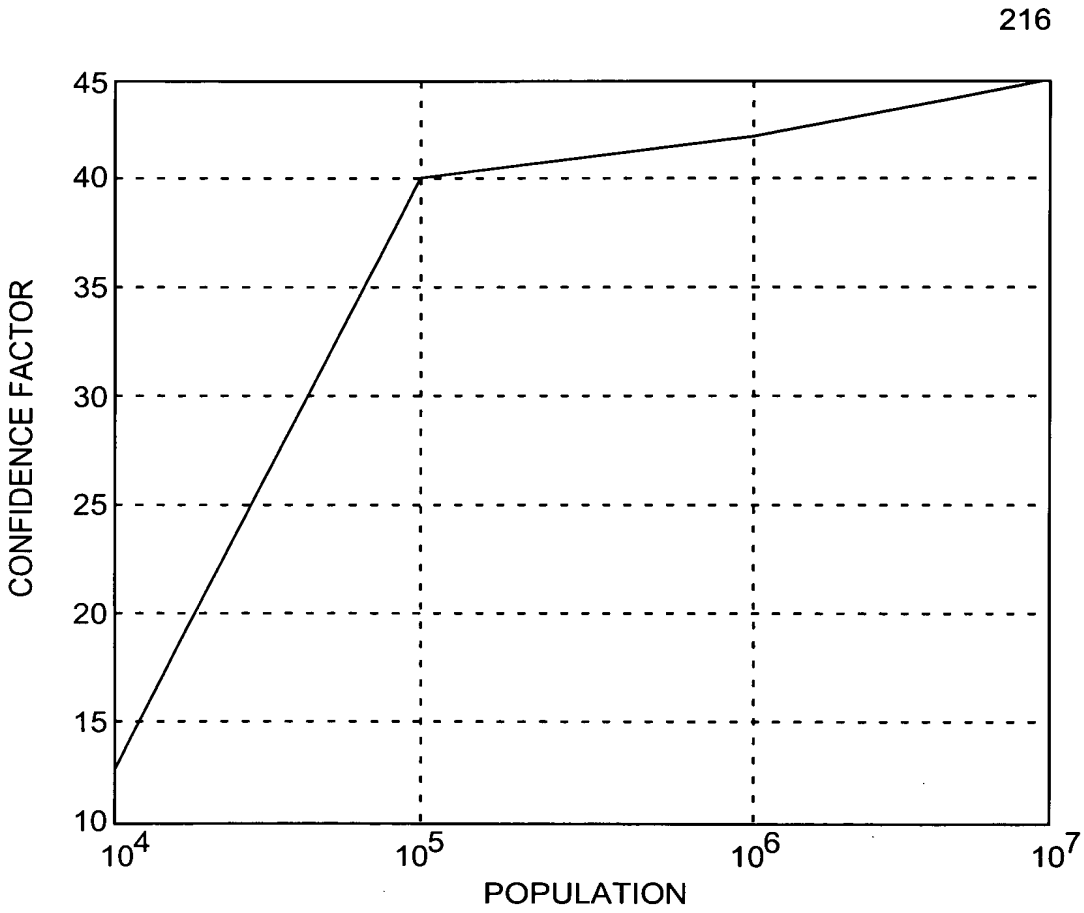
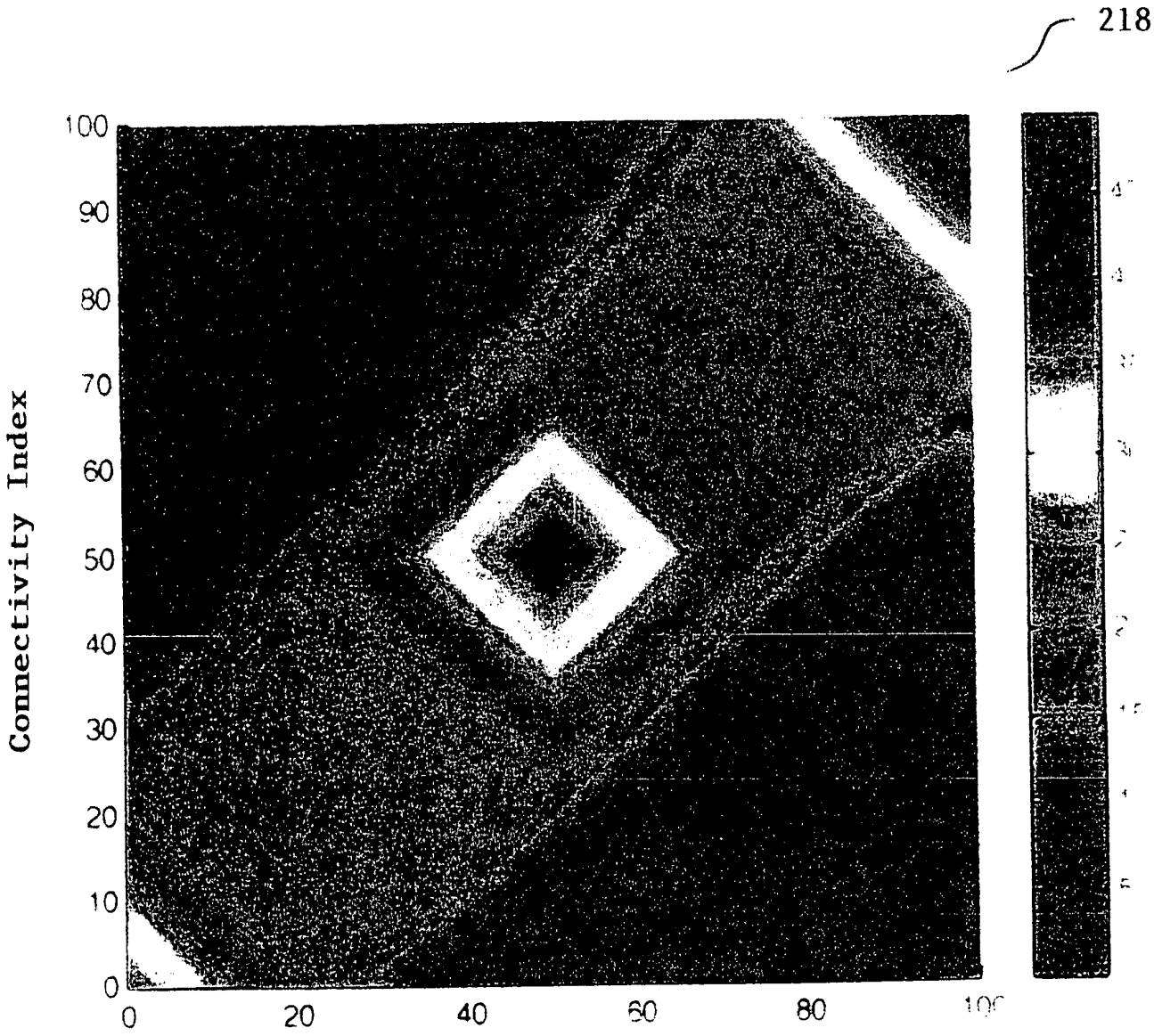

216
45
40
35
30
25
20
15
10
CONFIDENCE FACTOR
10^4
10^5
10^6
10^7
POPULATION

FIG. 14N



Means of Last Known Hop and Next Known Hop Connectivity Indices

FIG. 140

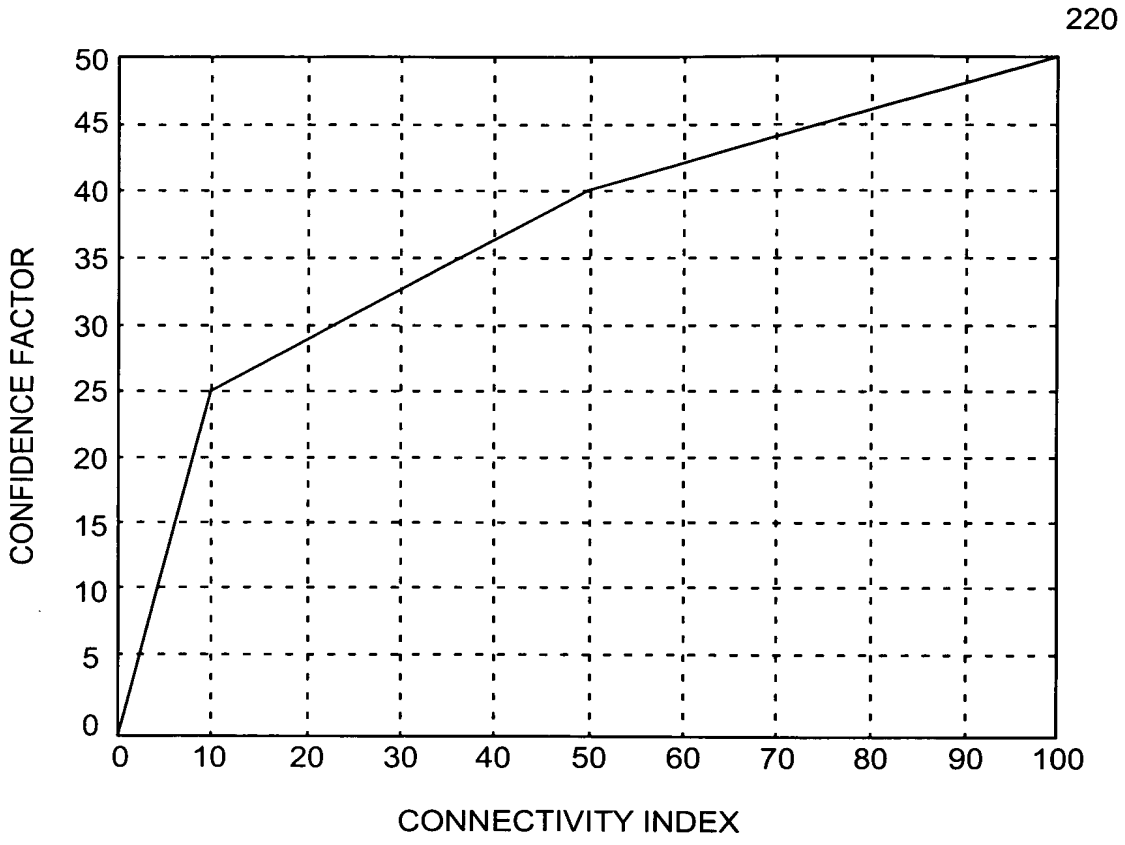

220
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
10
20
30
40
50
60
70
80
90
100
CONNECTIVITY INDEX

FIG. 14P

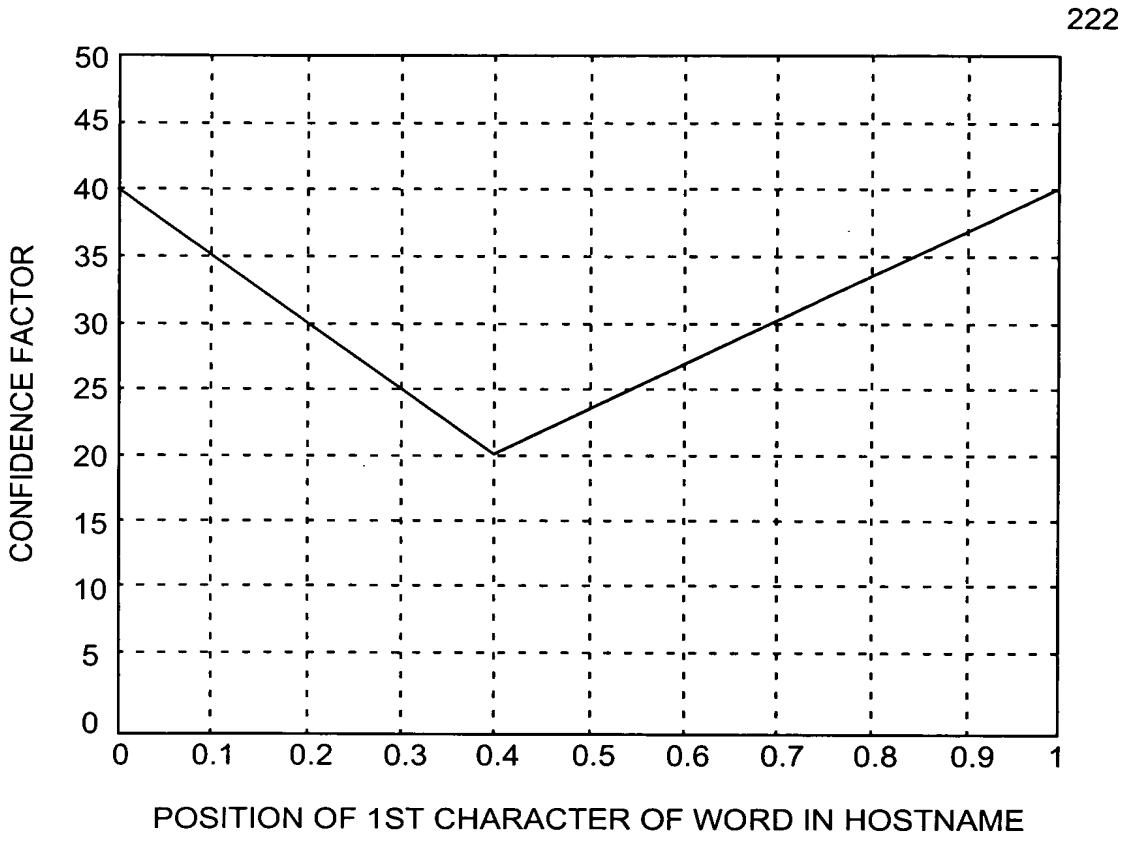

222
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
POSITION OF 1ST CHARACTER OF WORD IN HOSTNAME

FIG. 14Q

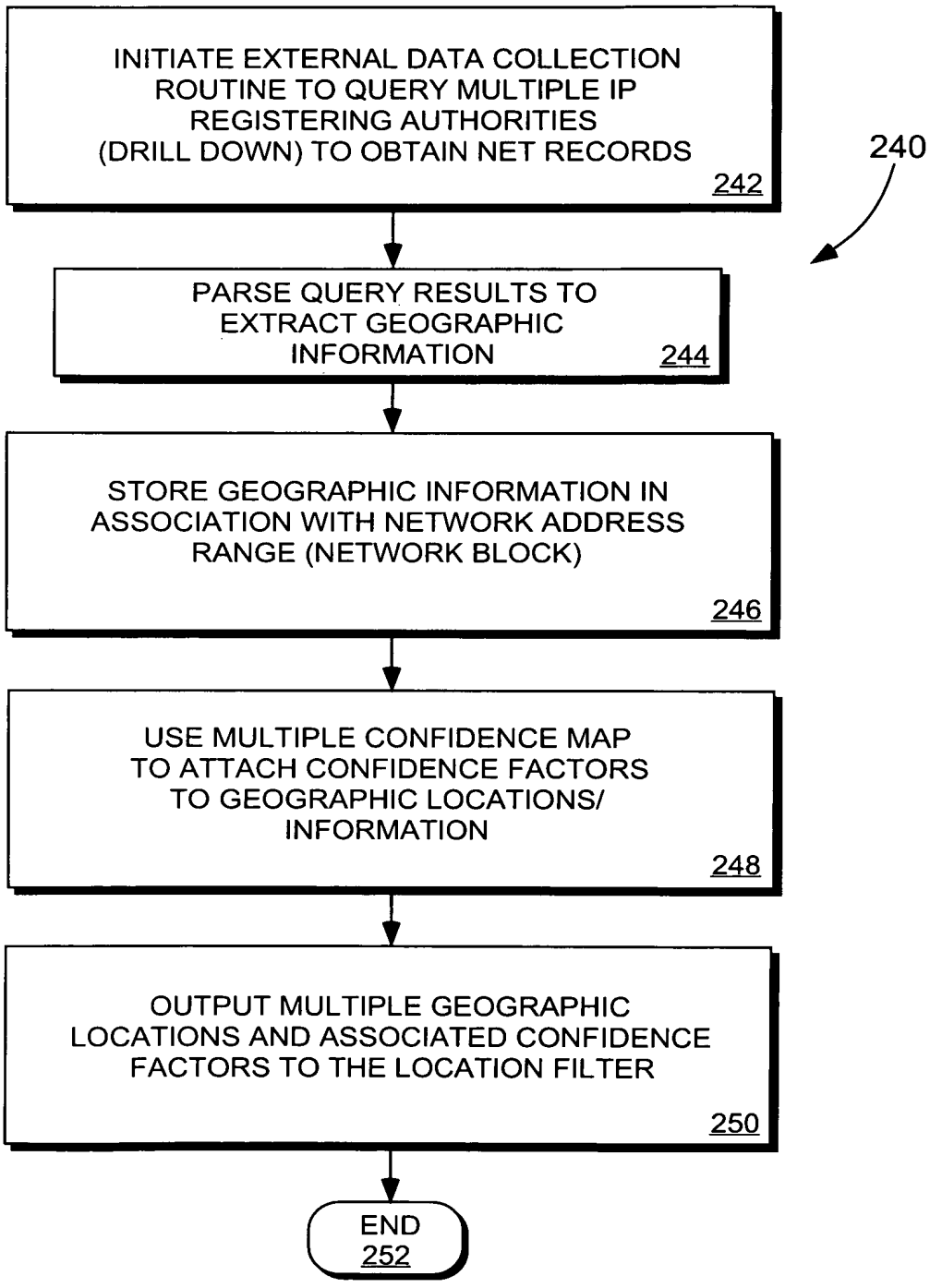

INITIATE EXTERNAL DATA COLLECTION ROUTINE TO QUERY MULTIPLE IP REGISTERING AUTHORITIES (DRILL DOWN) TO OBTAIN NET RECORDS
242
240
PARSE QUERY RESULTS TO EXTRACT GEOGRAPHIC INFORMATION
244
STORE GEOGRAPHIC INFORMATION IN ASSOCIATION WITH NETWORK ADDRESS RANGE (NETWORK BLOCK)
246
USE MULTIPLE CONFIDENCE MAP TO ATTACH CONFIDENCE FACTORS TO GEOGRAPHIC LOCATIONS/ INFORMATION
248
OUTPUT MULTIPLE GEOGRAPHIC LOCATIONS AND ASSOCIATED CONFIDENCE FACTORS TO THE LOCATION FILTER
250
END
252

FIG. 15

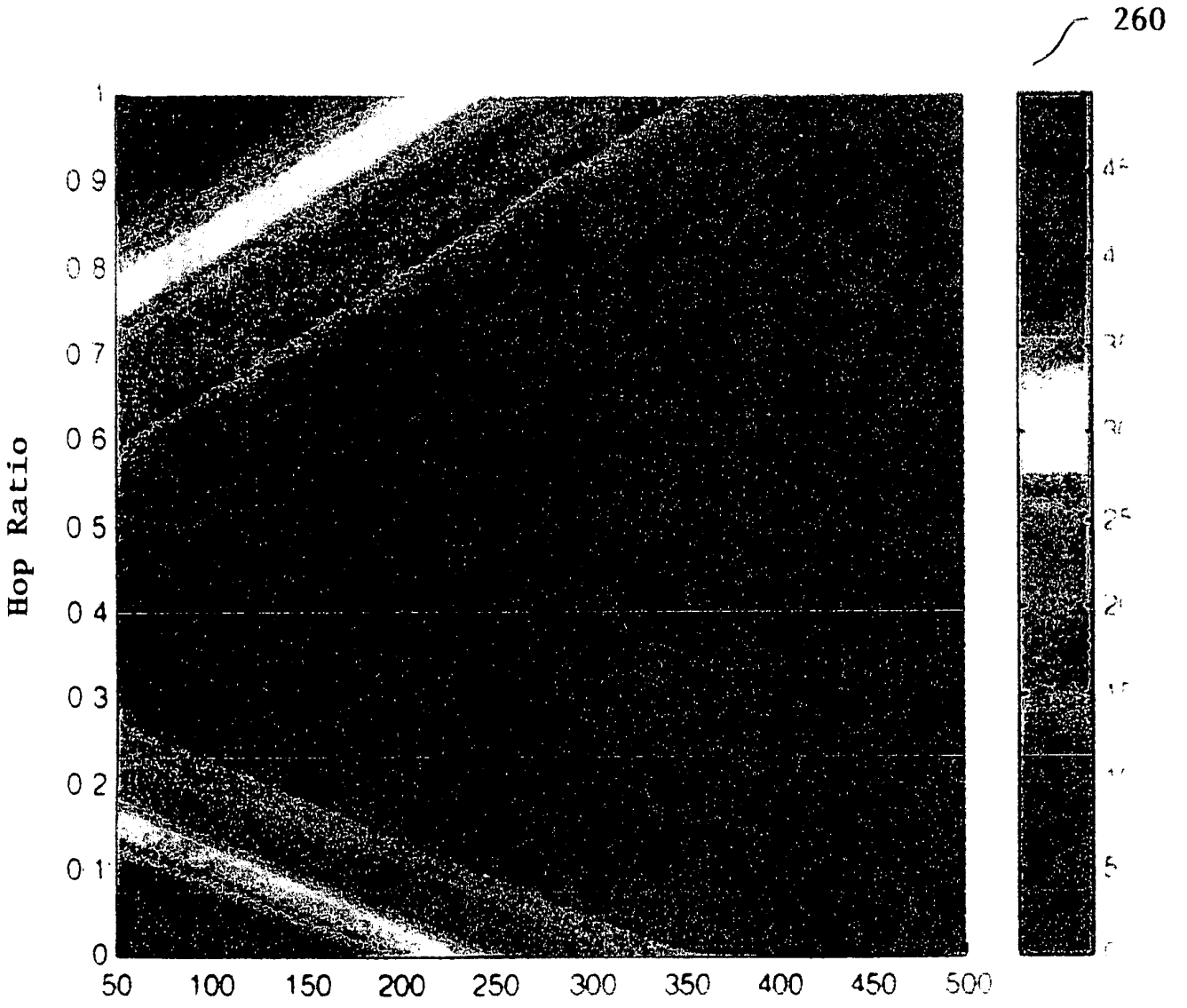

260
Hop Ratio
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
50
100
150
200
250
300
350
400
450
500
Distance in Miles Between Last Known Hop and Net

FIG. 16A

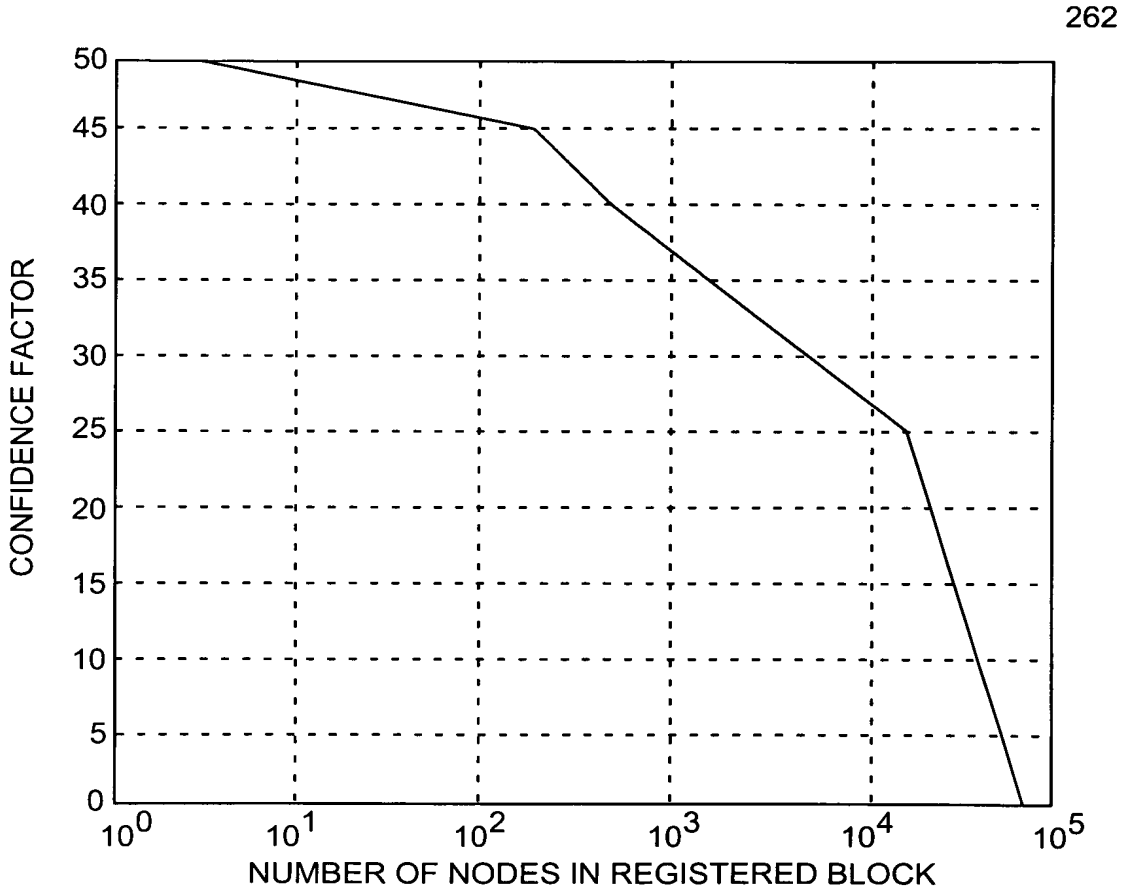

262
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
10^0
10^1
10^2
10^3
10^4
10^5
NUMBER OF NODES IN REGISTERED BLOCK

FIG. 16B

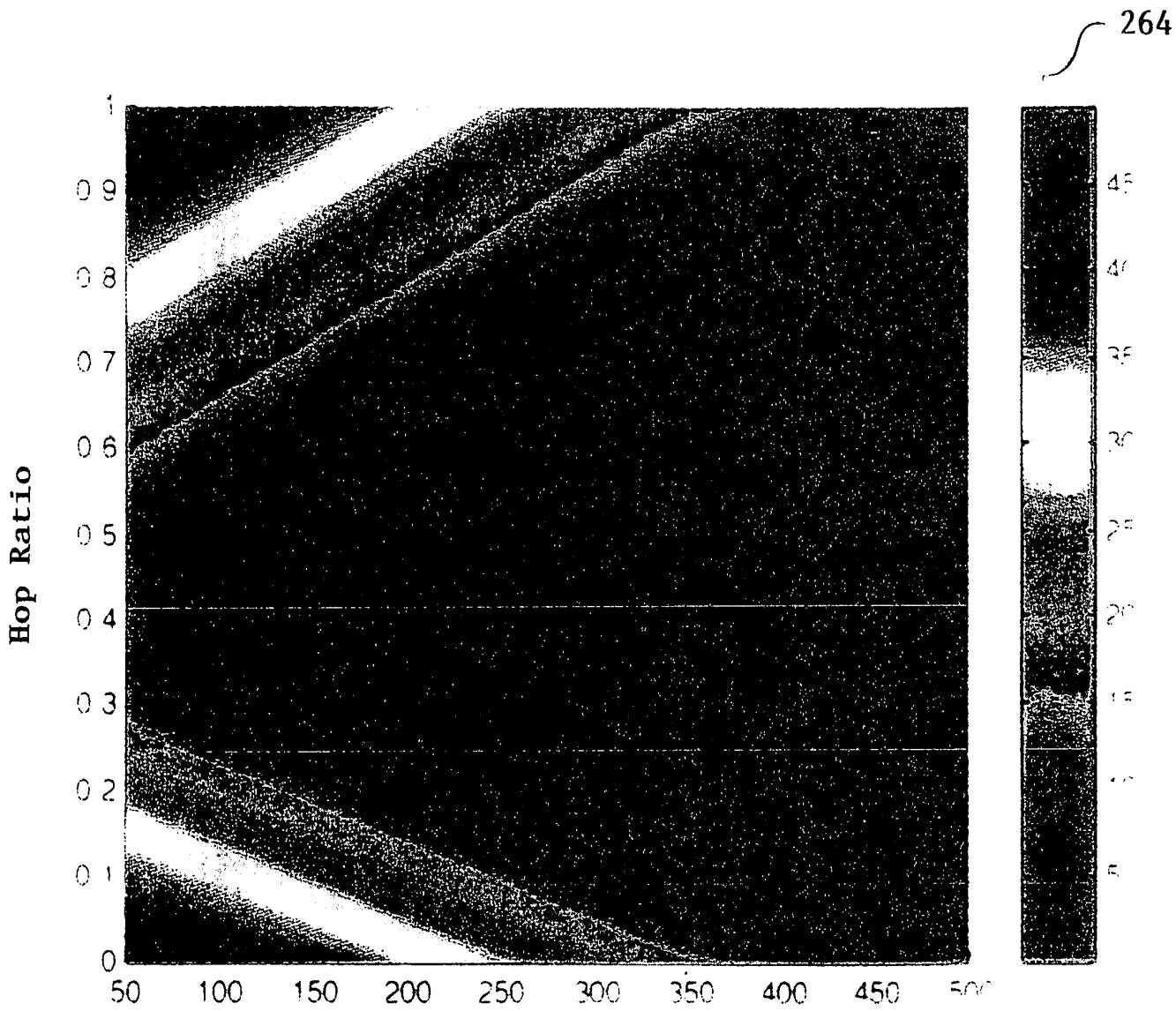

264
Hop Ratio
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
50
100
150
200
250
300
350
400
450
Distance in Miles Between Last Known Hop and Net

FIG. 16C

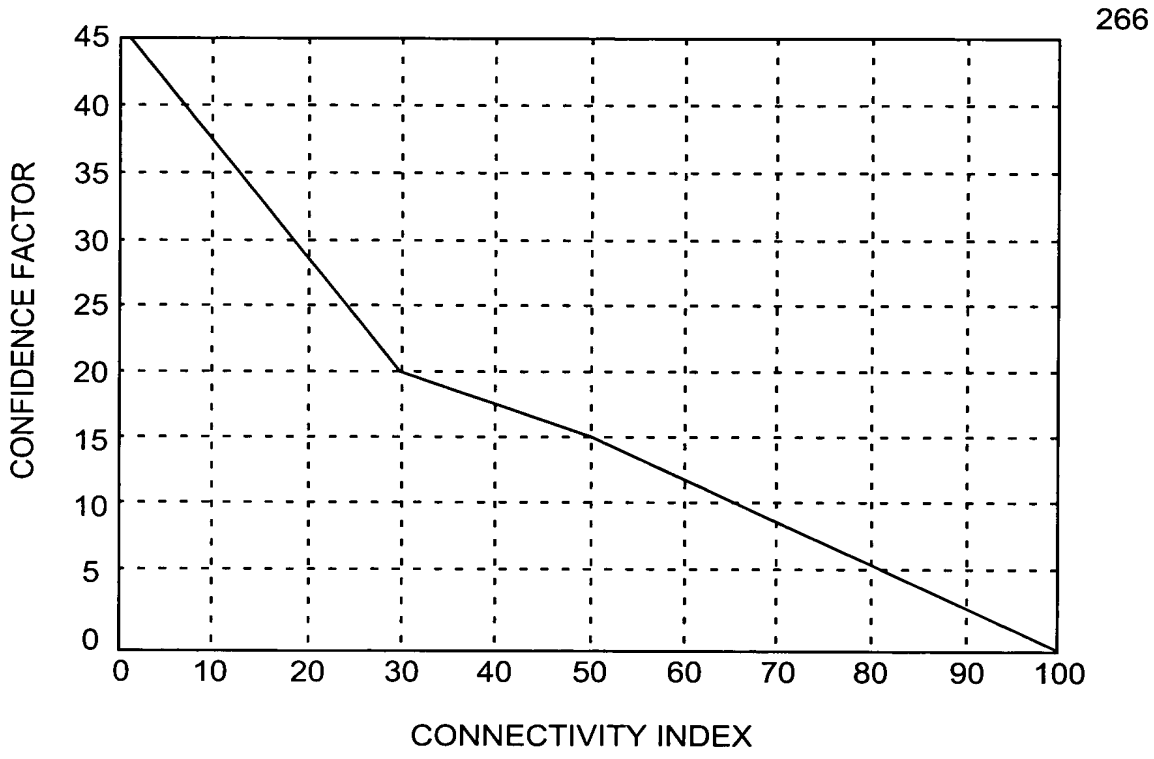

266
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
10
20
30
40
50
60
70
80
90
100
CONNECTIVITY INDEX

FIG. 16D

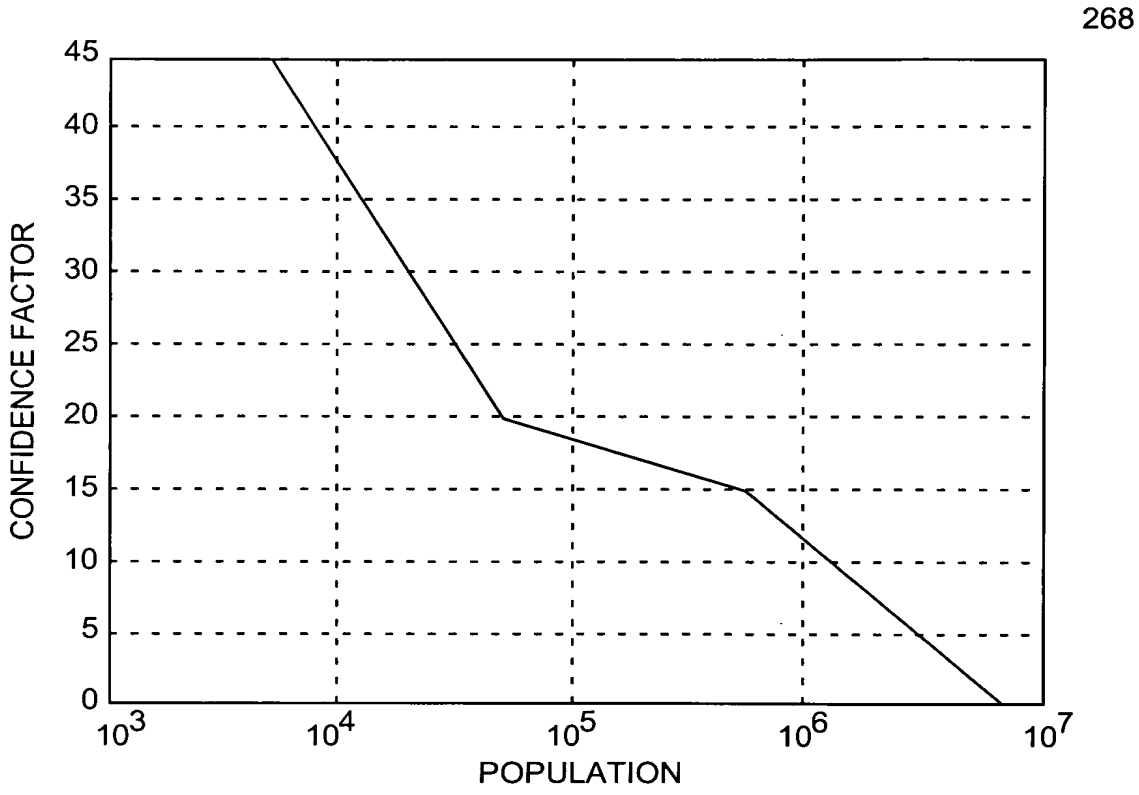

268
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
10^3
10^4
10^5
10^6
10^7
POPULATION

FIG. 16E

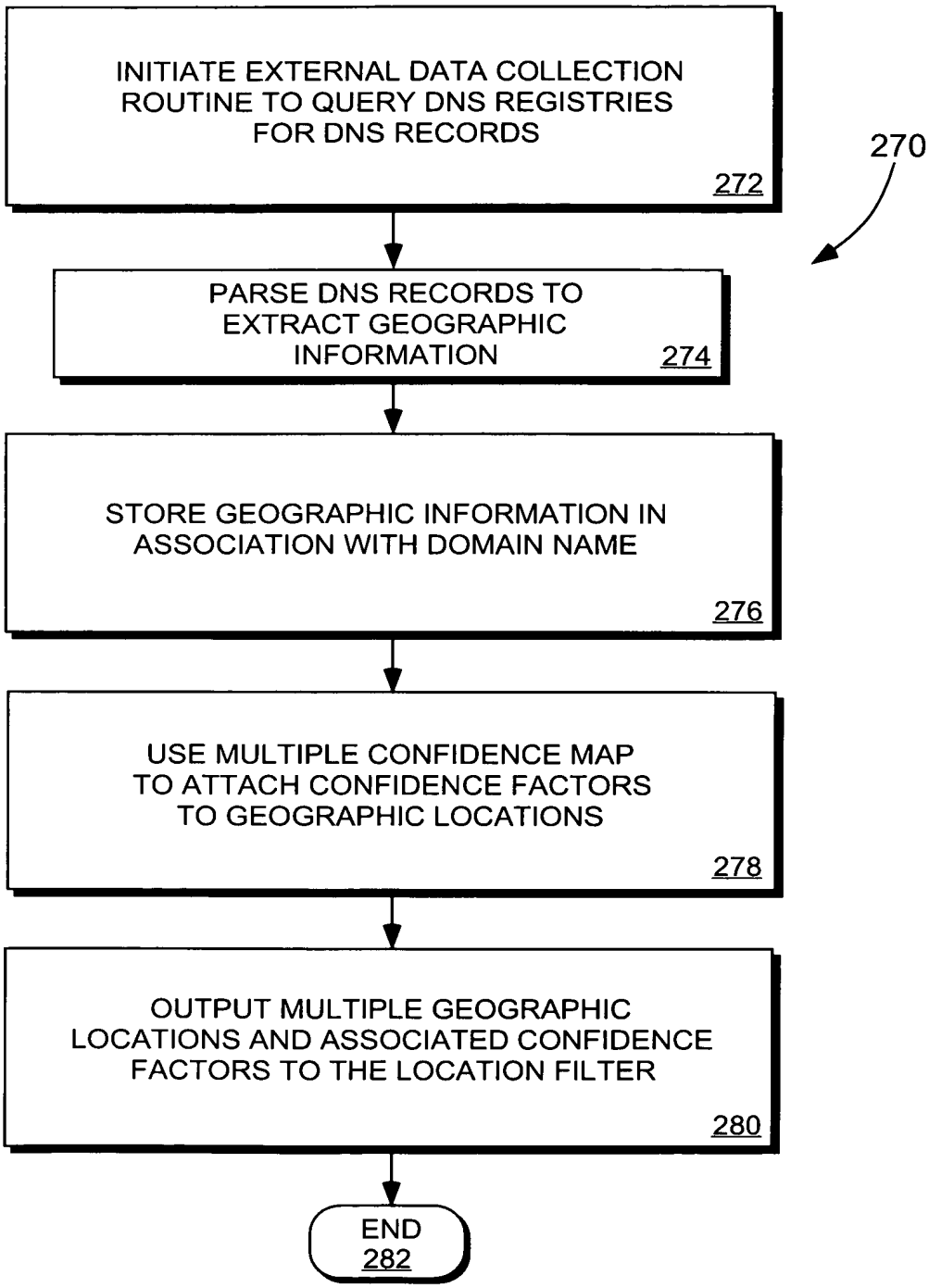

INITIATE EXTERNAL DATA COLLECTION ROUTINE TO QUERY DNS REGISTRIES FOR DNS RECORDS
272
270
PARSE DNS RECORDS TO EXTRACT GEOGRAPHIC INFORMATION
274
STORE GEOGRAPHIC INFORMATION IN ASSOCIATION WITH DOMAIN NAME
276
USE MULTIPLE CONFIDENCE MAP TO ATTACH CONFIDENCE FACTORS TO GEOGRAPHIC LOCATIONS
278
OUTPUT MULTIPLE GEOGRAPHIC LOCATIONS AND ASSOCIATED CONFIDENCE FACTORS TO THE LOCATION FILTER
280
END
282

FIG. 17

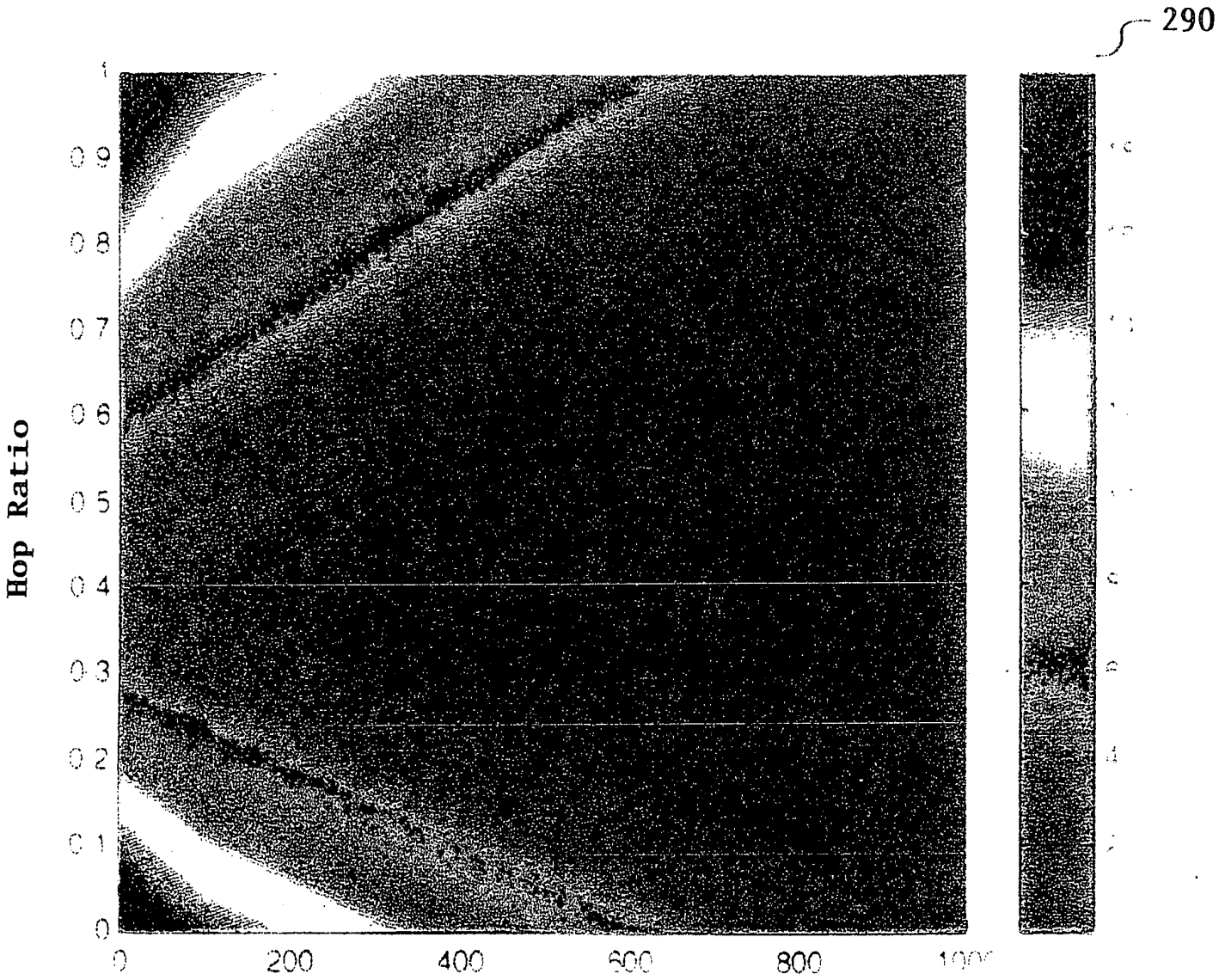

290
Hop Ratio
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
200
400
800
Distance in Miles Between Last Known Hop and DNS

FIG. 18A

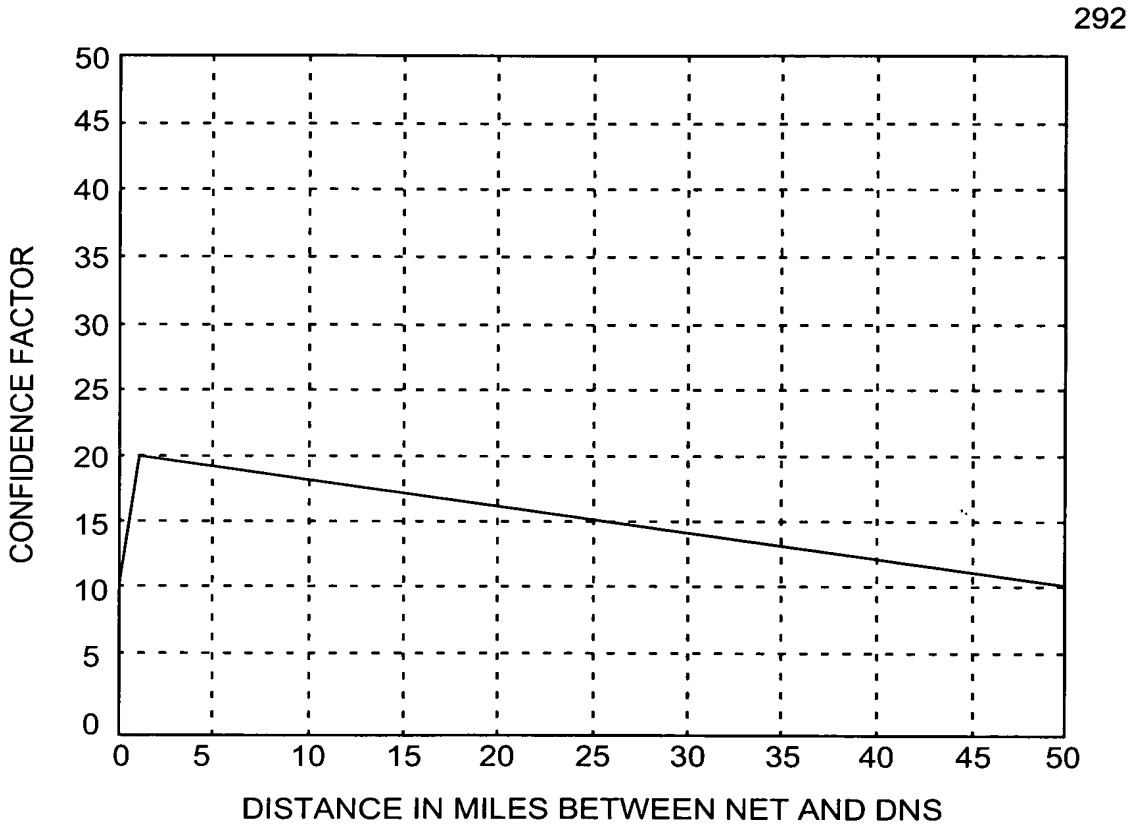

292
CONFIDENCE FACTOR
50
45
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
35
40
45
50
DISTANCE IN MILES BETWEEN NET AND DNS

FIG. 18B

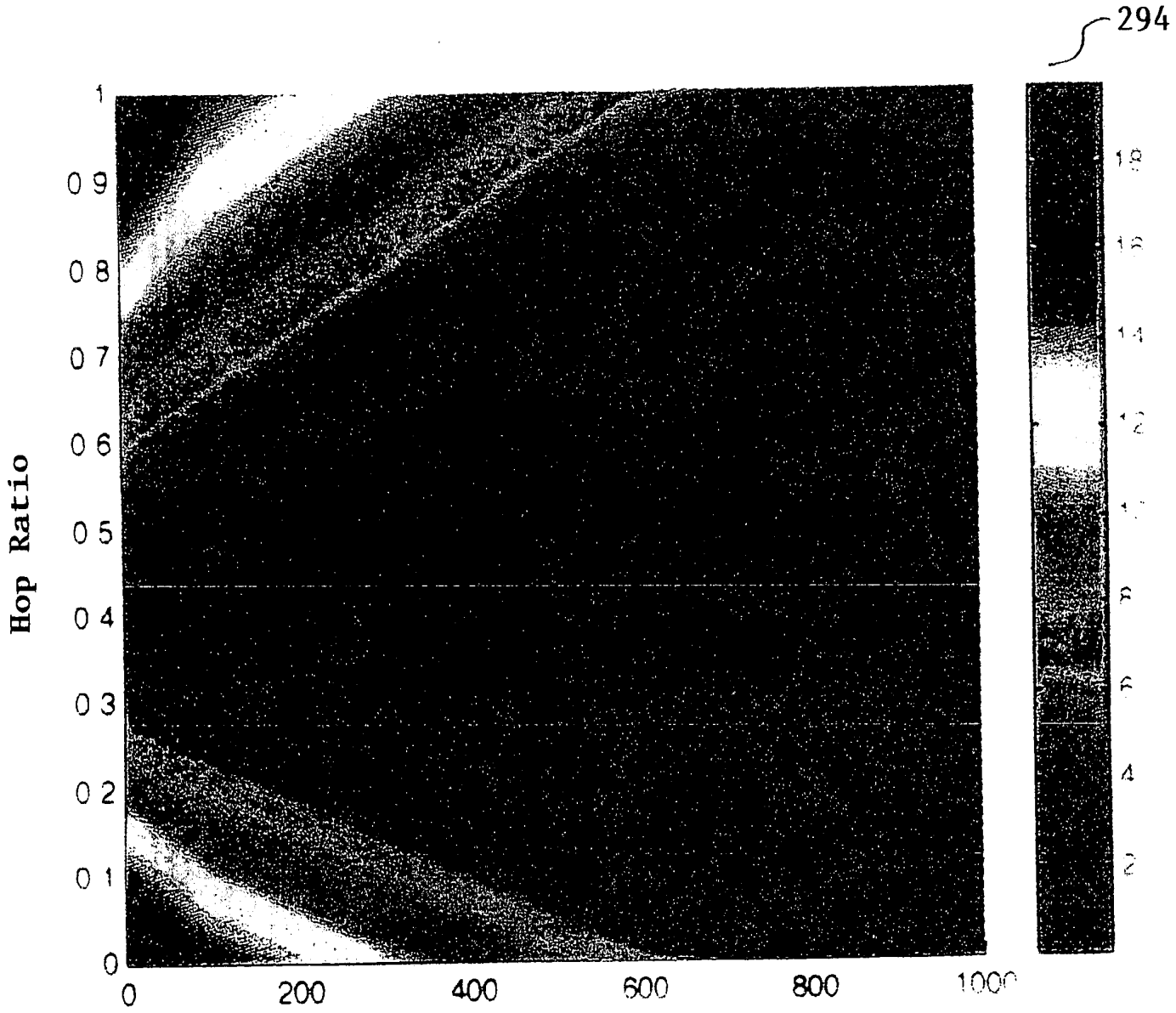

294
Hop Ratio
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
0
200
400
600
800
1000
Distance in Miles Between Next Known Hop and DNS

FIG. 18C

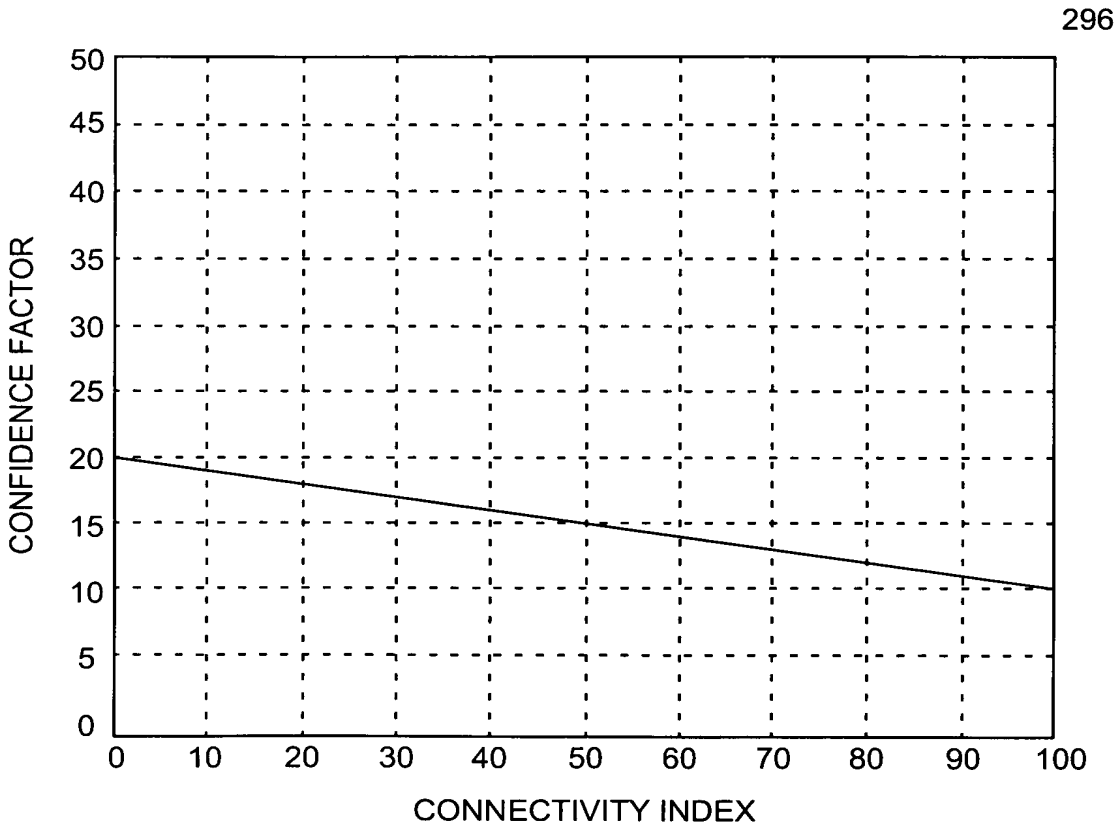

296
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
10
20
30
40
50
60
70
80
90
100
CONNECTIVITY INDEX

FIG. 18D

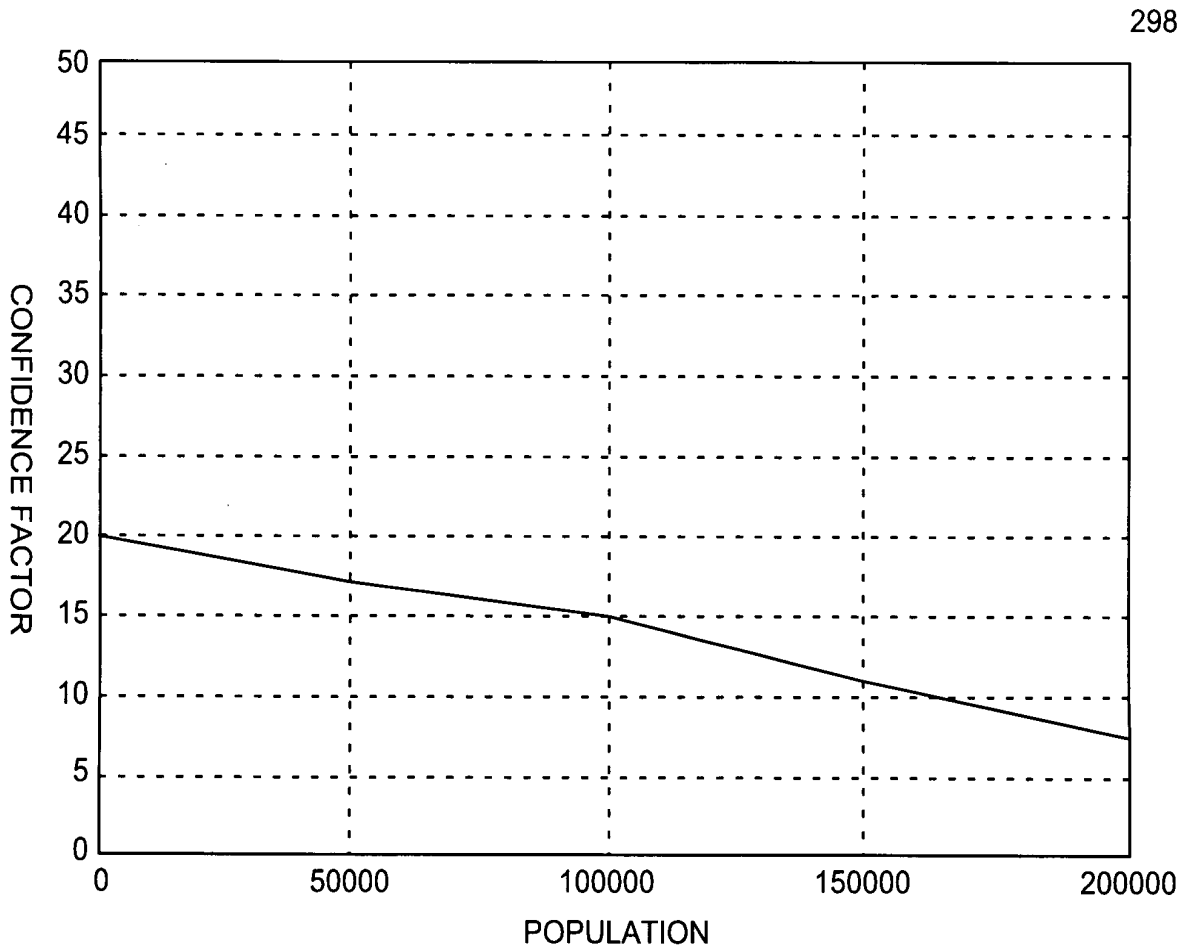

298
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
50000
100000
150000
200000
POPULATION

FIG. 18E

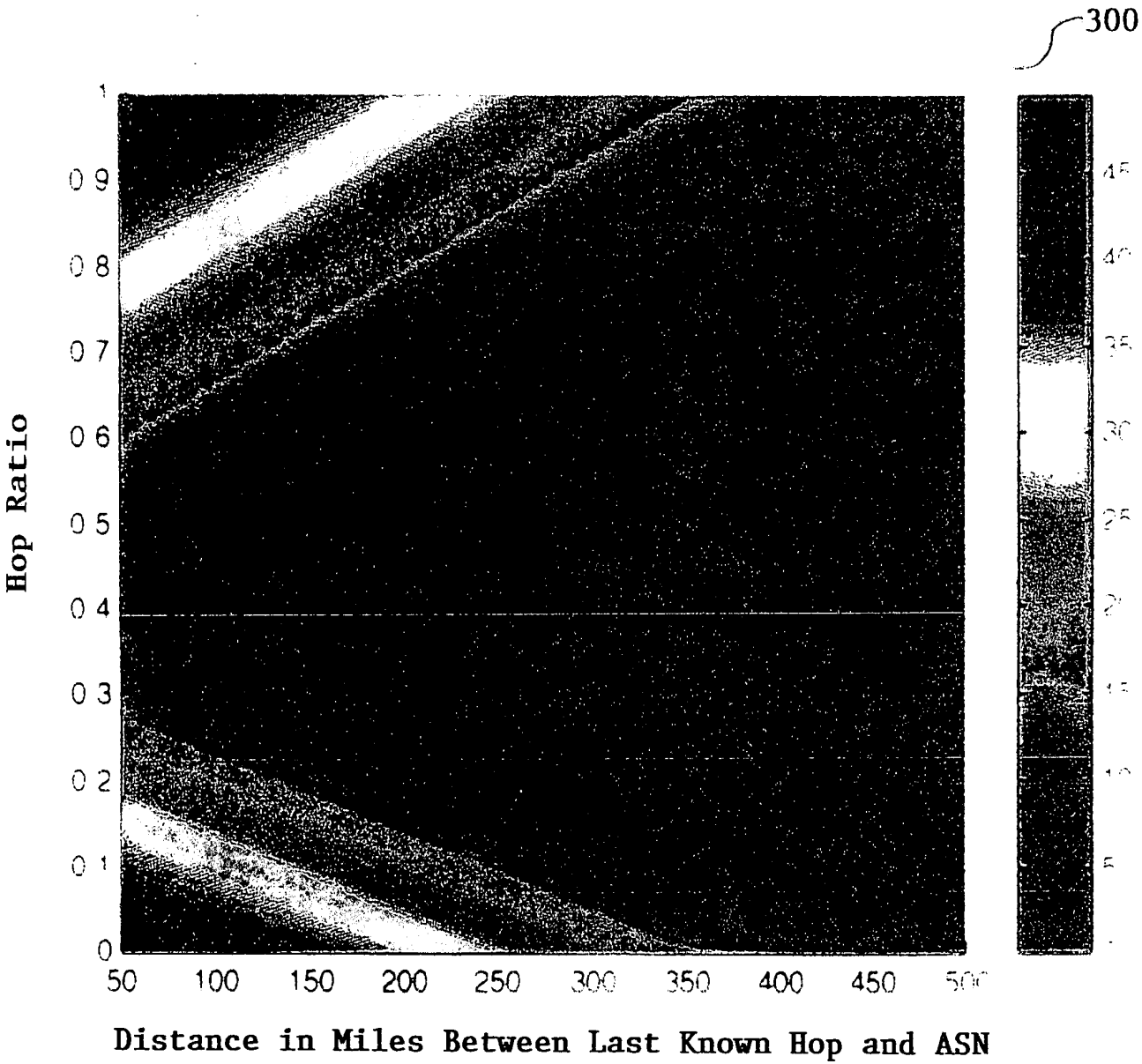

300
Hop Ratio
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
50
100
150
200
250
300
350
400
450
Distance in Miles Between Last Known Hop and ASN

FIG. 19A

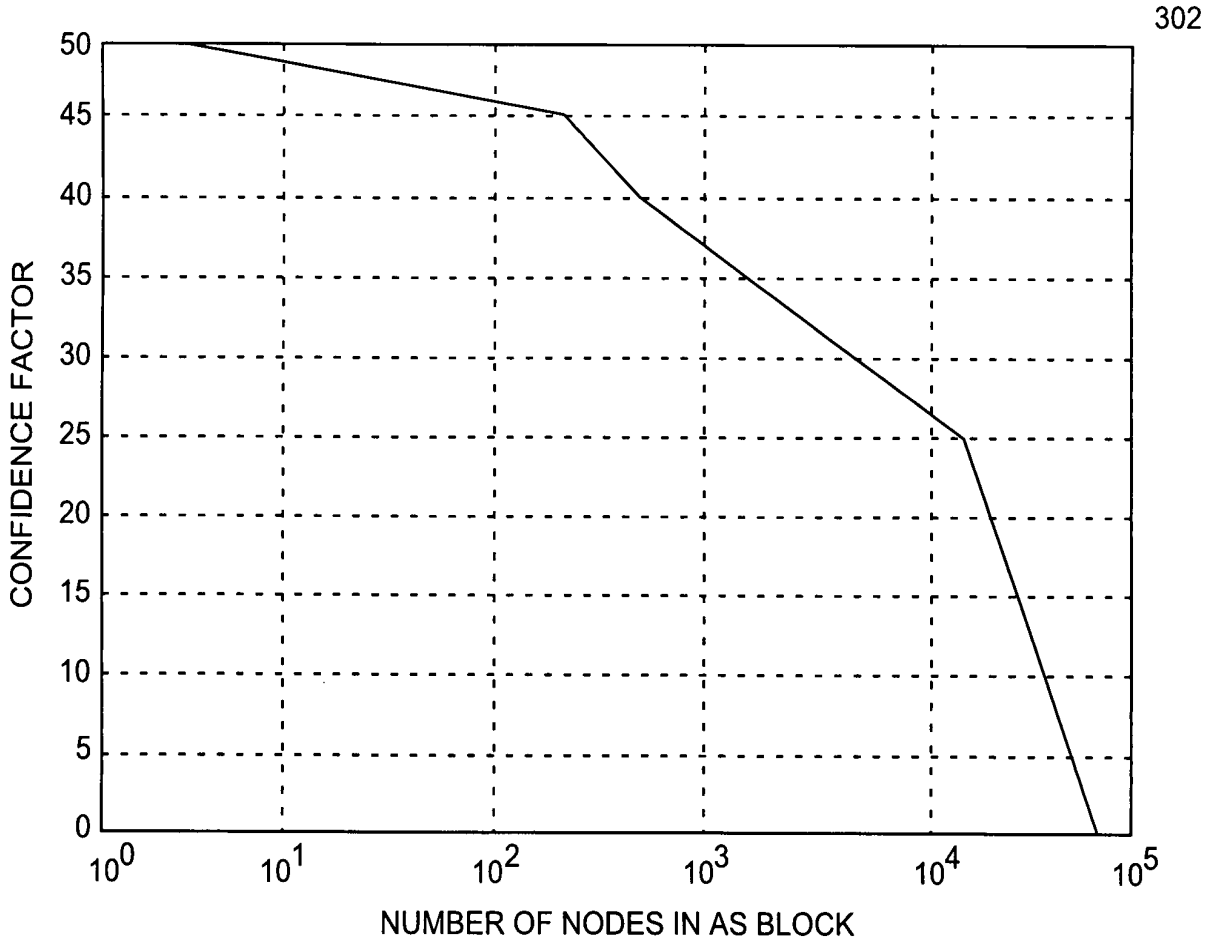

302
50
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
10^0
10^1
10^2
10^3
10^4
10^5
NUMBER OF NODES IN AS BLOCK

FIG. 19B

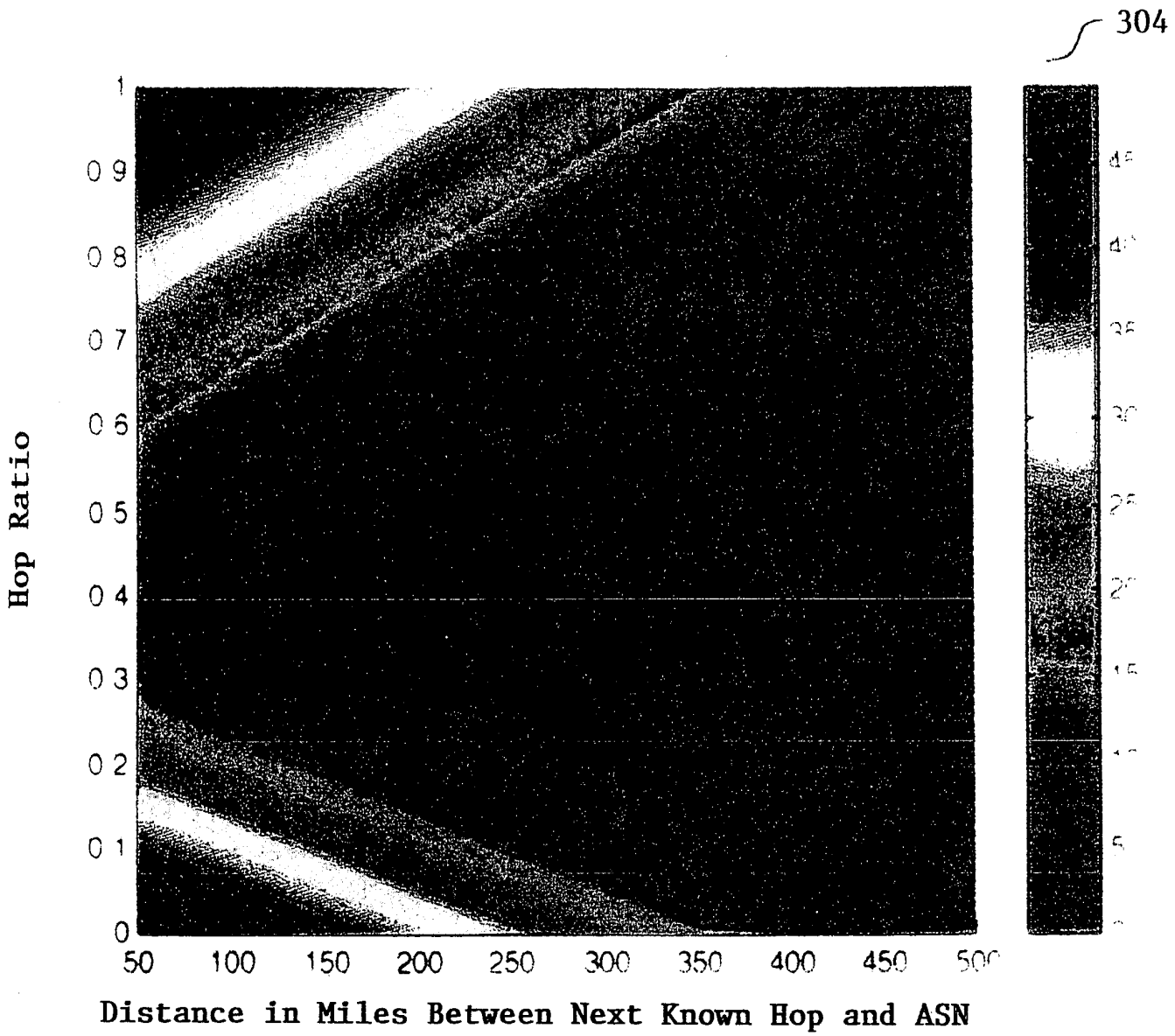

304
Hop Ratio
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
50
100
150
200
250
300
350
400
450
Distance in Miles Between Next Known Hop and ASN

FIG. 19C

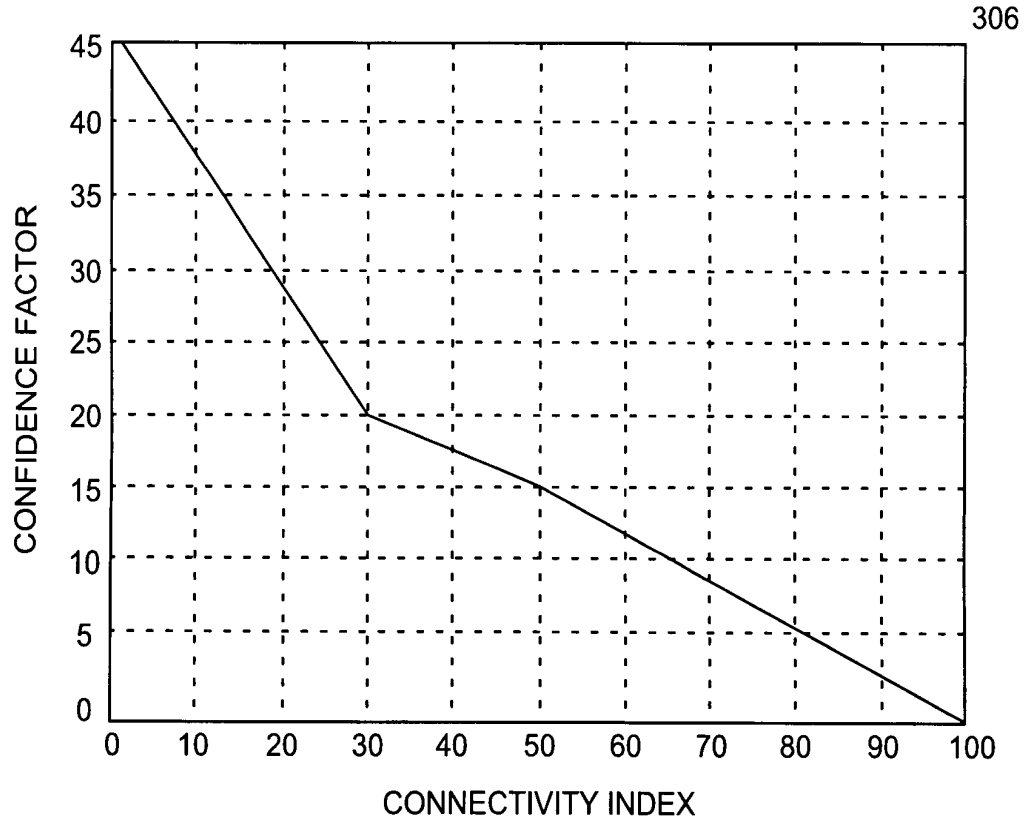

306
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
0
10
20
30
40
50
60
70
80
90
100
CONNECTIVITY INDEX

FIG. 19D

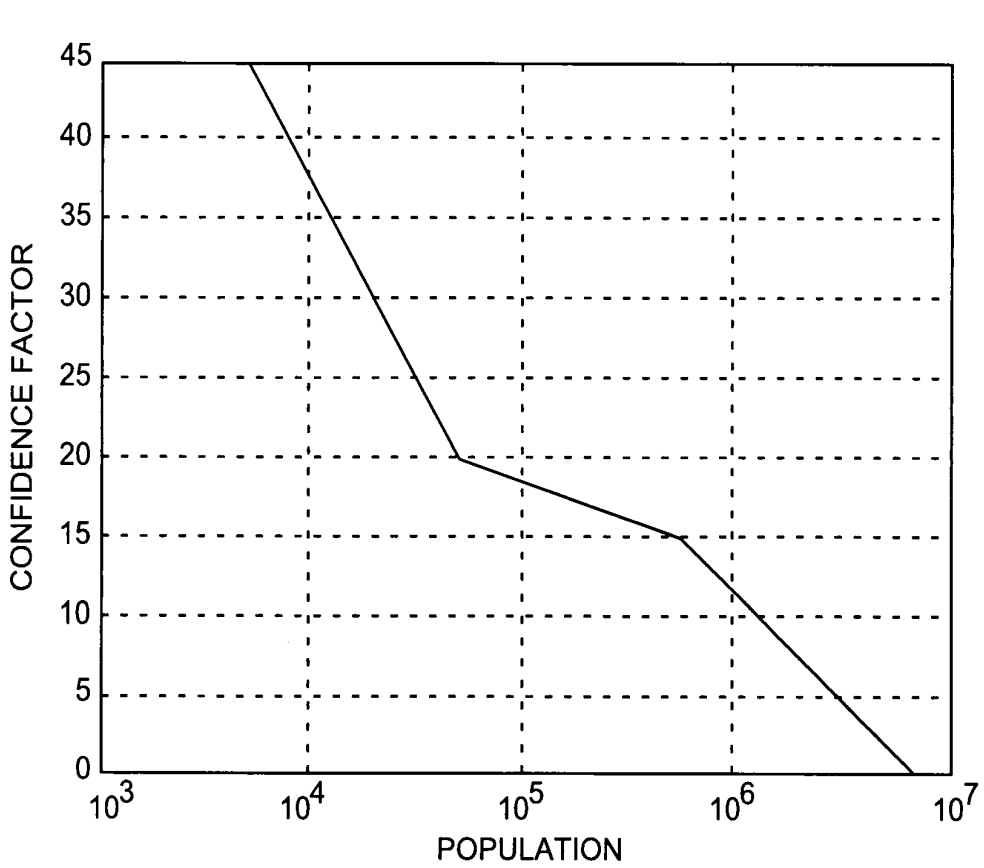

308
45
40
35
30
25
20
15
10
5
0
CONFIDENCE FACTOR
10^3
10^4
10^5
10^6
10^7
POPULATION

FIG. 19E

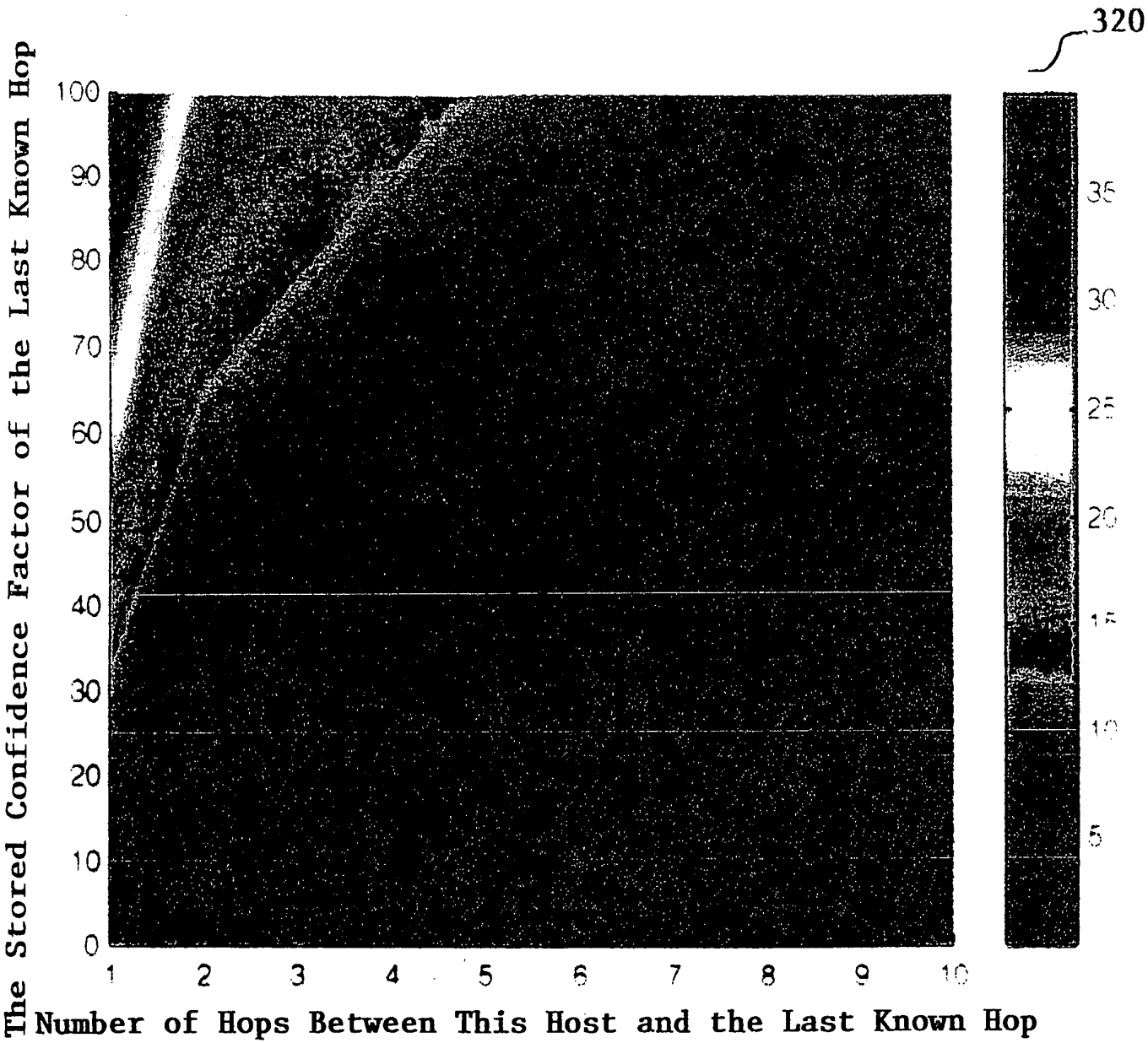

320
The Stored Confidence Factor of the Last Known Hop
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Number of Hops Between This Host and the Last Known Hop
35
30
25
20
15
10
5

FIG. 20A

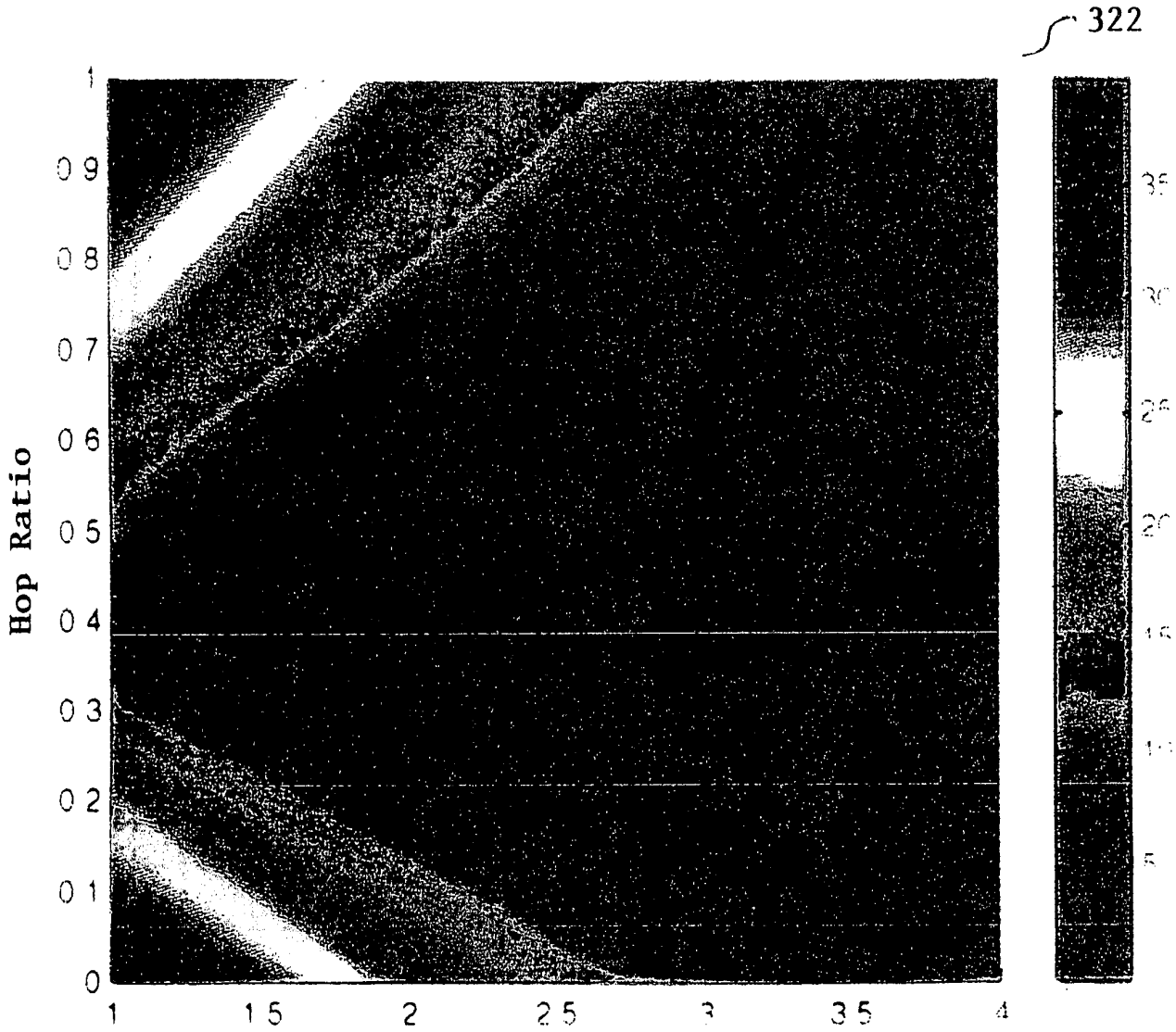

322
Hop Ratio
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
1
1 5
2
2 5
3
3 5
4
Number of Hops Between This Host and the Last Known Hop

FIG. 20B

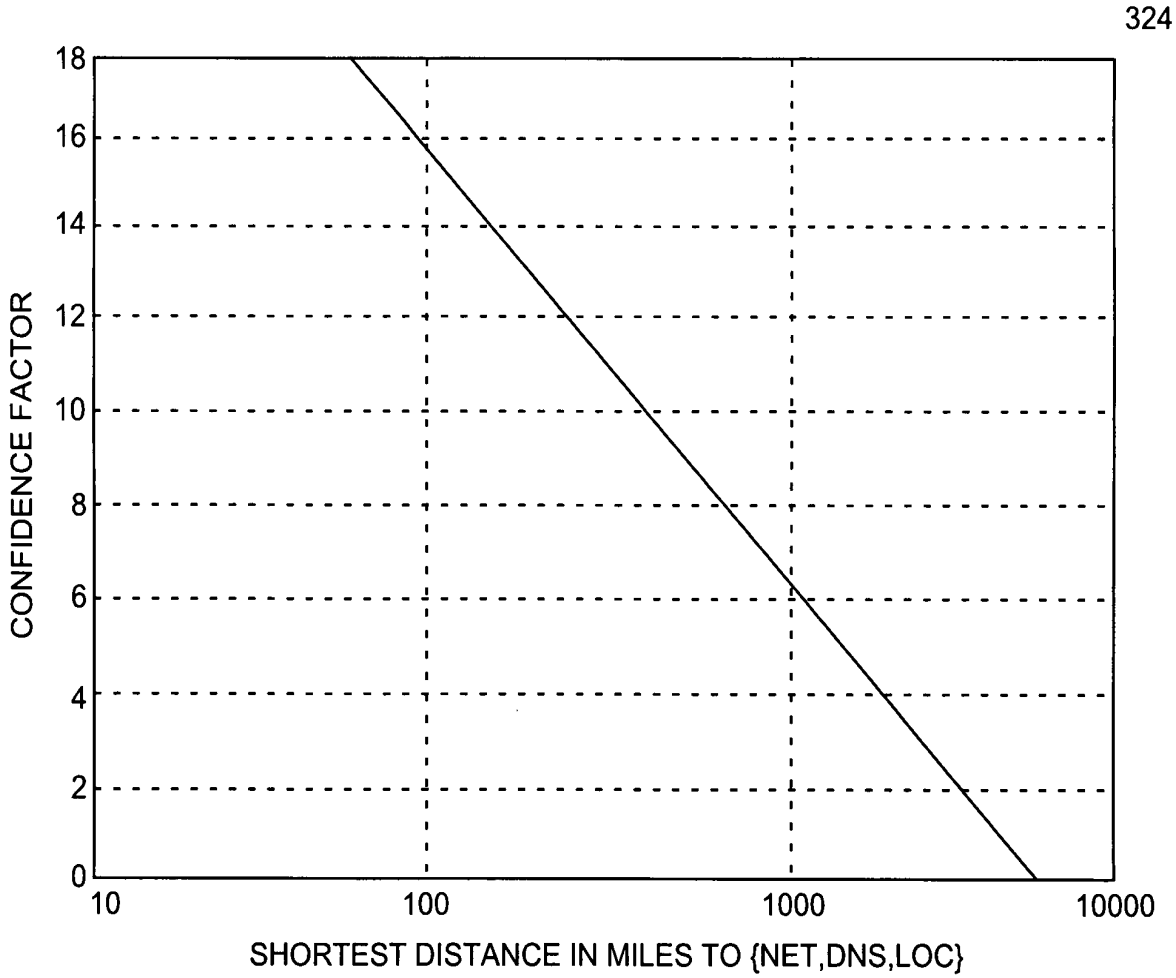

324
18
16
14
12
10
8
6
4
2
0
CONFIDENCE FACTOR
10
100
1000
10000
SHORTEST DISTANCE IN MILES TO {NET,DNS,LOC}

FIG. 20C

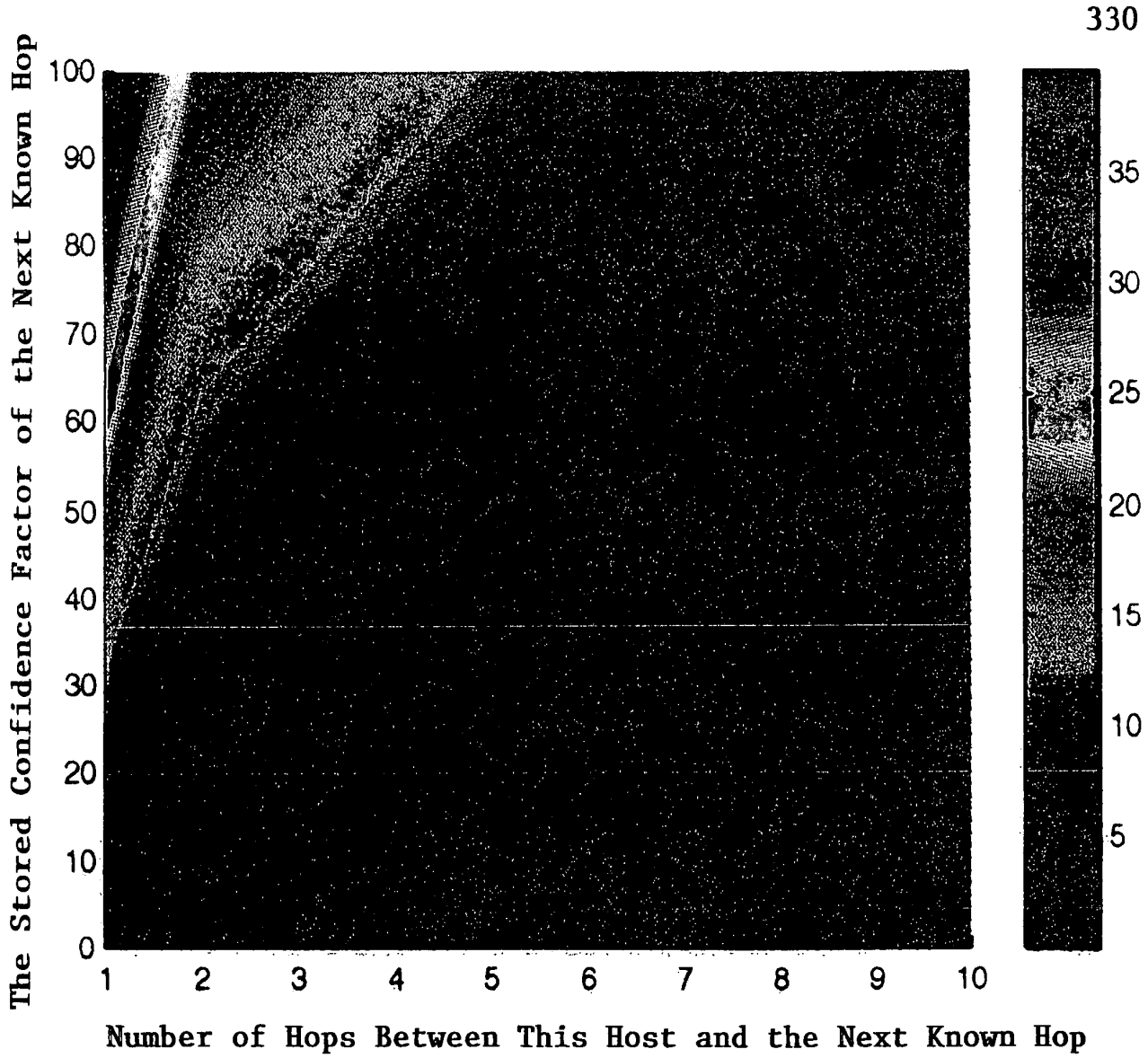

330
The Stored Confidence Factor of the Next Known Hop
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Number of Hops Between This Host and the Next Known Hop
35
30
25
20
15
10
5

FIG. 21A

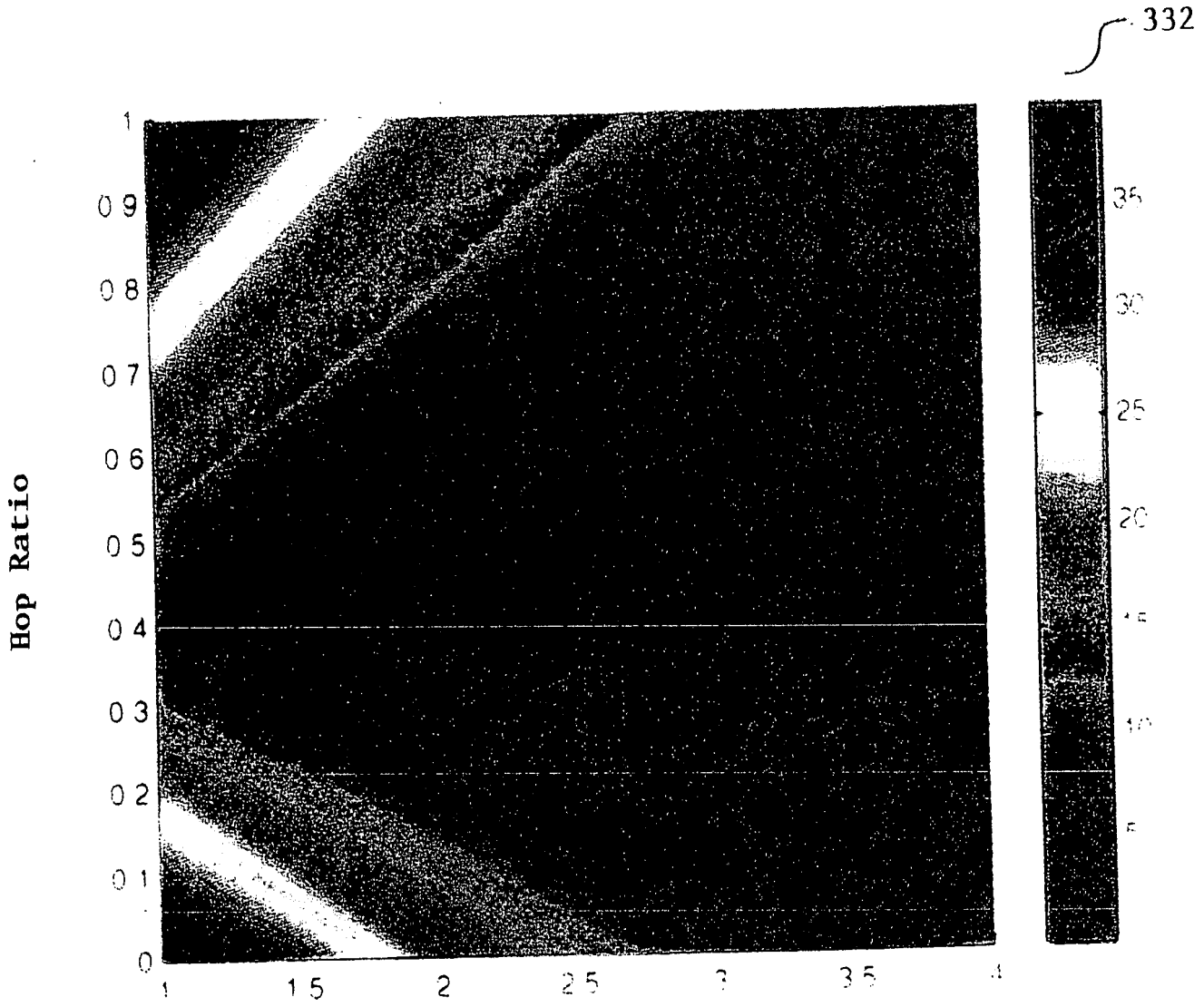

332
Hop Ratio
1
0 9
0 8
0 7
0 6
0 5
0 4
0 3
0 2
0 1
0
1
1 5
2
2 5
3
3 5
Number of Hops Between This Host and the Next Known Hop

FIG. 21B

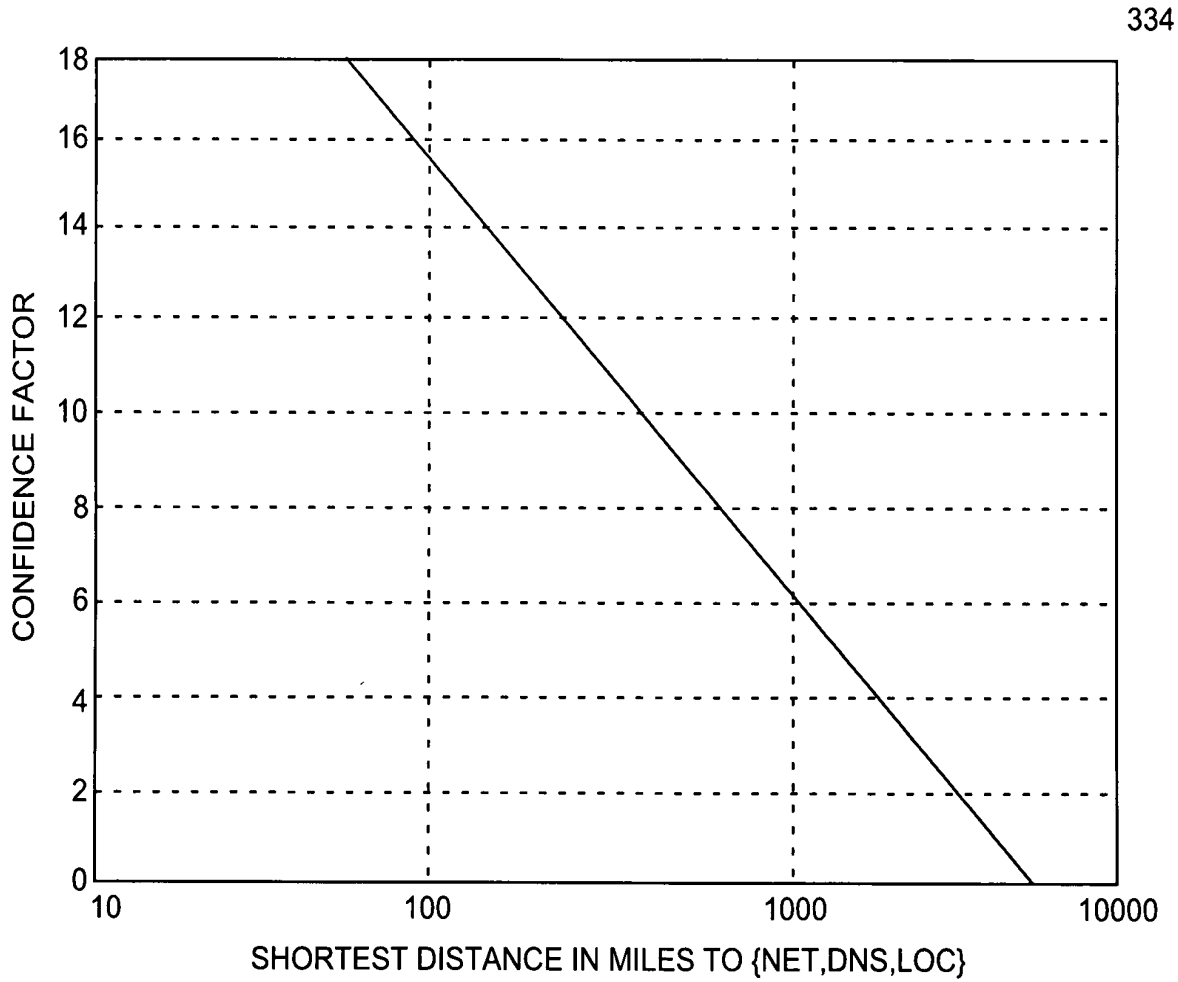

334
18
16
14
12
10
8
6
4
2
0
CONFIDENCE FACTOR
10
100
1000
10000
SHORTEST DISTANCE IN MILES TO {NET,DNS,LOC}

FIG. 21C

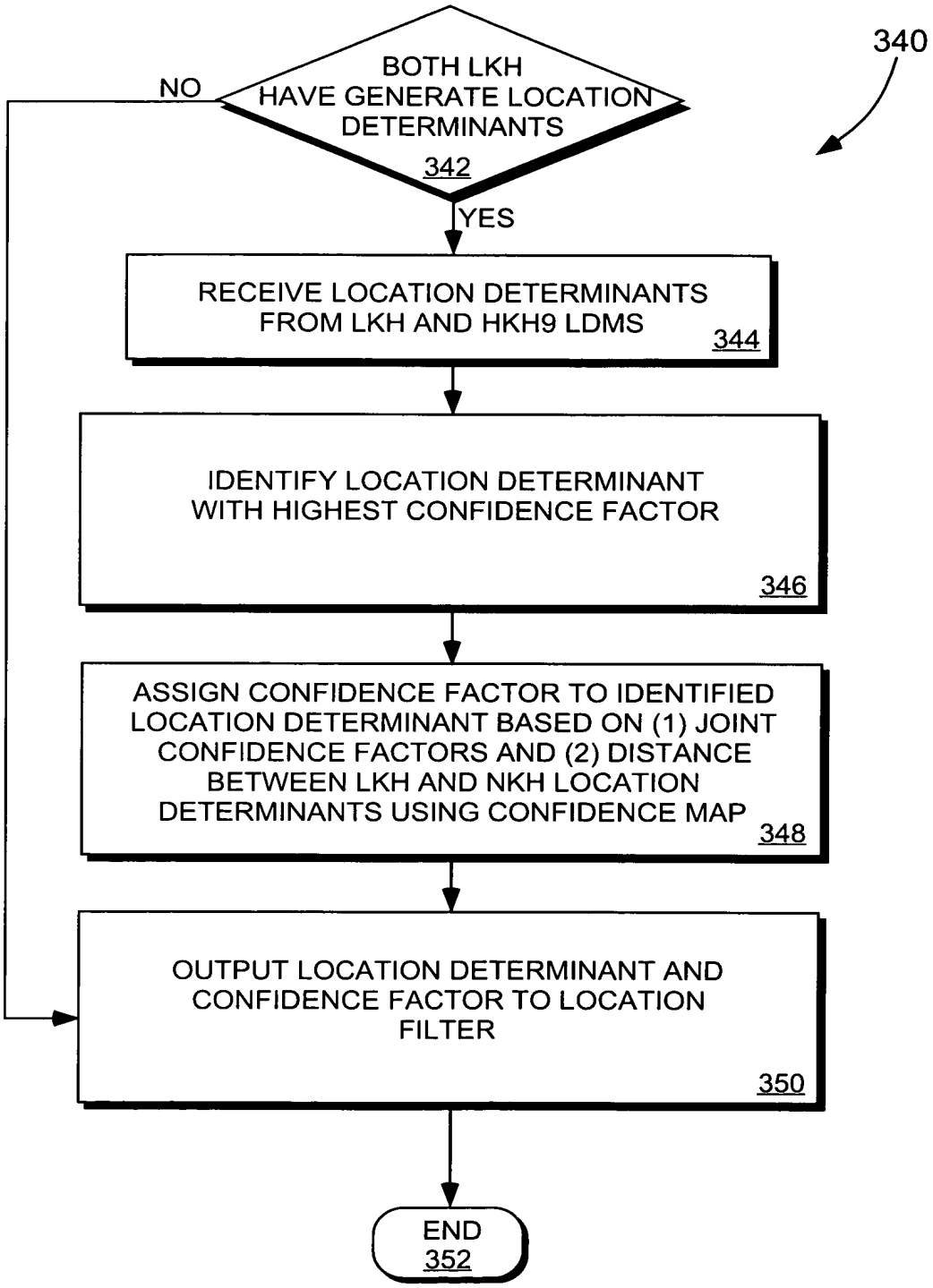

340
BOTH LKH HAVE GENERATE LOCATION DETERMINANTS
342
NO
YES
RECEIVE LOCATION DETERMINANTS FROM LKH AND HKH9 LDMS
344
IDENTIFY LOCATION DETERMINANT WITH HIGHEST CONFIDENCE FACTOR
346
ASSIGN CONFIDENCE FACTOR TO IDENTIFIED LOCATION DETERMINANT BASED ON (1) JOINT CONFIDENCE FACTORS AND (2) DISTANCE BETWEEN LKH AND NKH LOCATION DETERMINANTS USING CONFIDENCE MAP
348
OUTPUT LOCATION DETERMINANT AND CONFIDENCE FACTOR TO LOCATION FILTER
350
END
352

FIG. 22

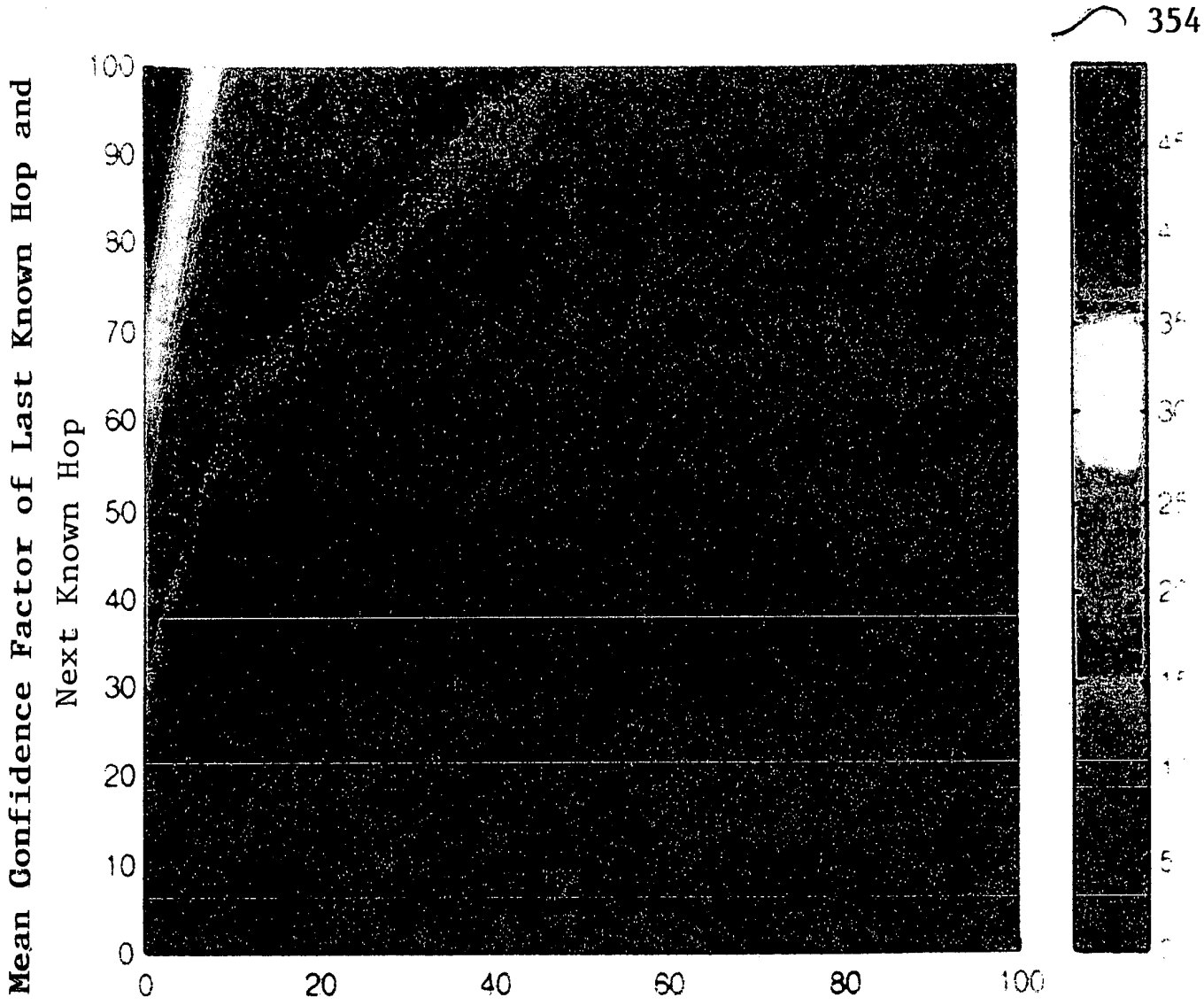

354
Mean Confidence Factor of Last Known Hop and Next Known Hop
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
Distance in Miles Between Last Known Hop and Next Known Hop

FIG. 23

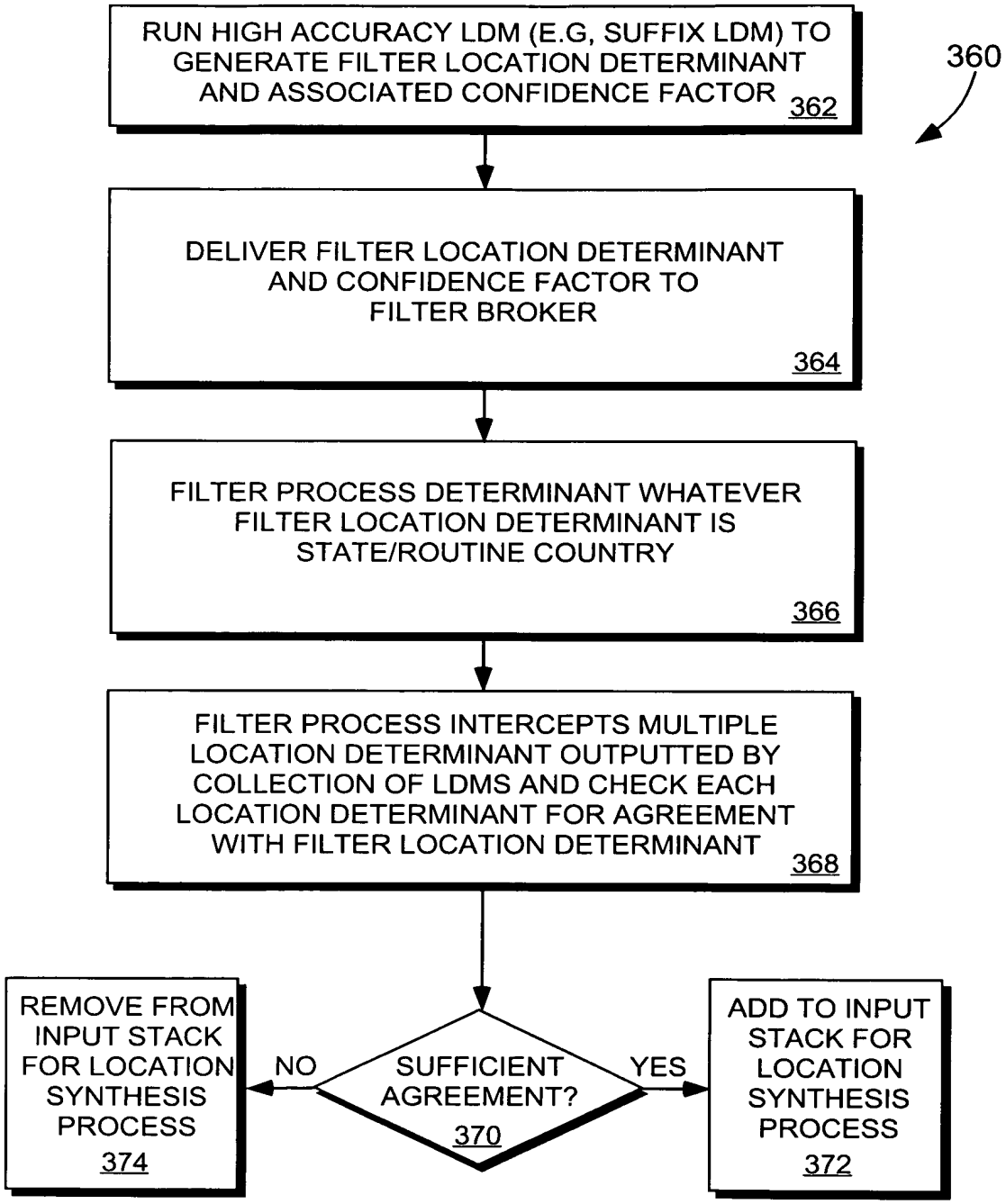

RUN HIGH ACCURACY LDM (E.G, SUFFIX LDM) TO GENERATE FILTER LOCATION DETERMINANT AND ASSOCIATED CONFIDENCE FACTOR
362
360
DELIVER FILTER LOCATION DETERMINANT AND CONFIDENCE FACTOR TO FILTER BROKER
364
FILTER PROCESS DETERMINANT WHATEVER FILTER LOCATION DETERMINANT IS STATE/ROUTINE COUNTRY
366
FILTER PROCESS INTERCEPTS MULTIPLE LOCATION DETERMINANT OUTPUTTED BY COLLECTION OF LDMS AND CHECK EACH LOCATION DETERMINANT FOR AGREEMENT WITH FILTER LOCATION DETERMINANT
368
SUFFICIENT AGREEMENT?
370
NO
YES
REMOVE FROM INPUT STACK FOR LOCATION SYNTHESIS PROCESS
374
ADD TO INPUT STACK FOR LOCATION SYNTHESIS PROCESS
372

FIG. 24

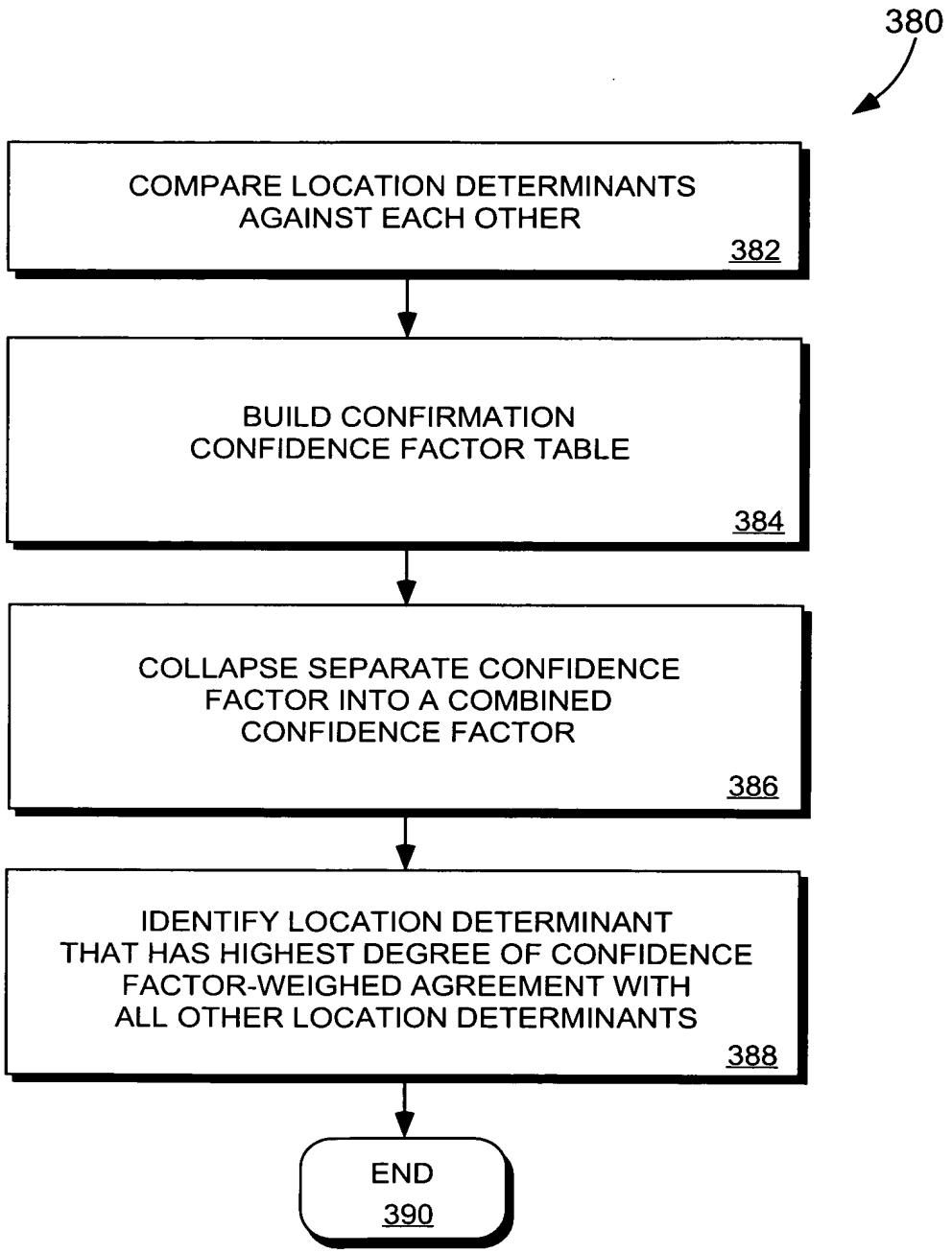

380
COMPARE LOCATION DETERMINANTS AGAINST EACH OTHER
382
BUILD CONFIRMATION CONFIDENCE FACTOR TABLE
384
COLLAPSE SEPARATE CONFIDENCE FACTOR INTO A COMBINED CONFIDENCE FACTOR
386
IDENTIFY LOCATION DETERMINANT THAT HAS HIGHEST DEGREE OF CONFIDENCE FACTOR-WEIGHED AGREEMENT WITH ALL OTHER LOCATION DETERMINANTS
388
END
390

FIG. 25

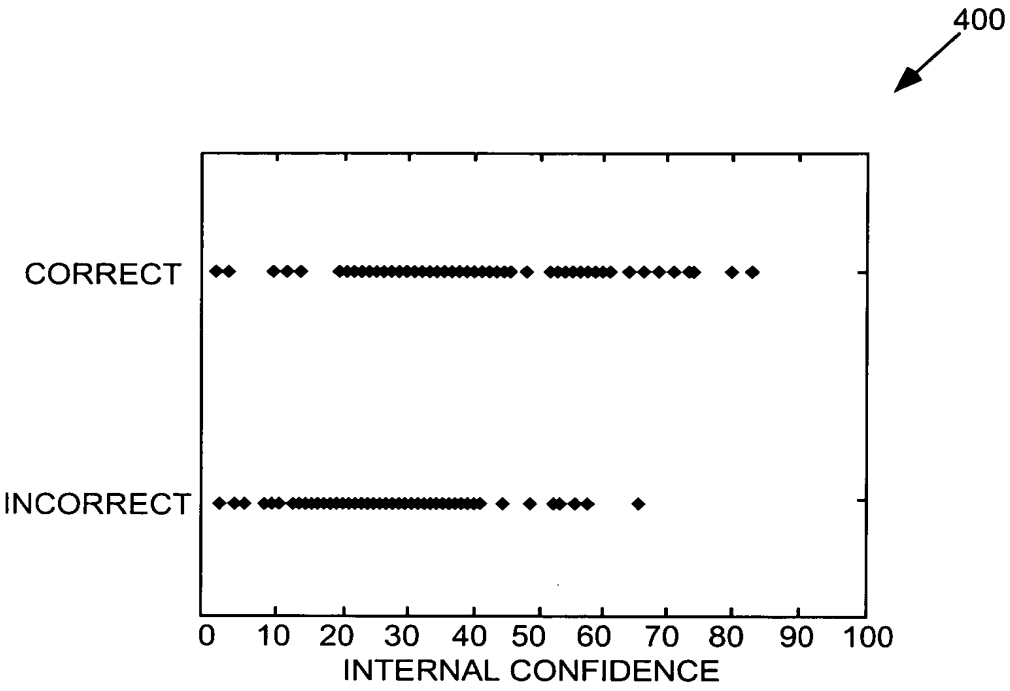

400
CORRECT
INCORRECT
0 10 20 30 40 50 60 70 80 90 100
INTERNAL CONFIDENCE

FIG. 26

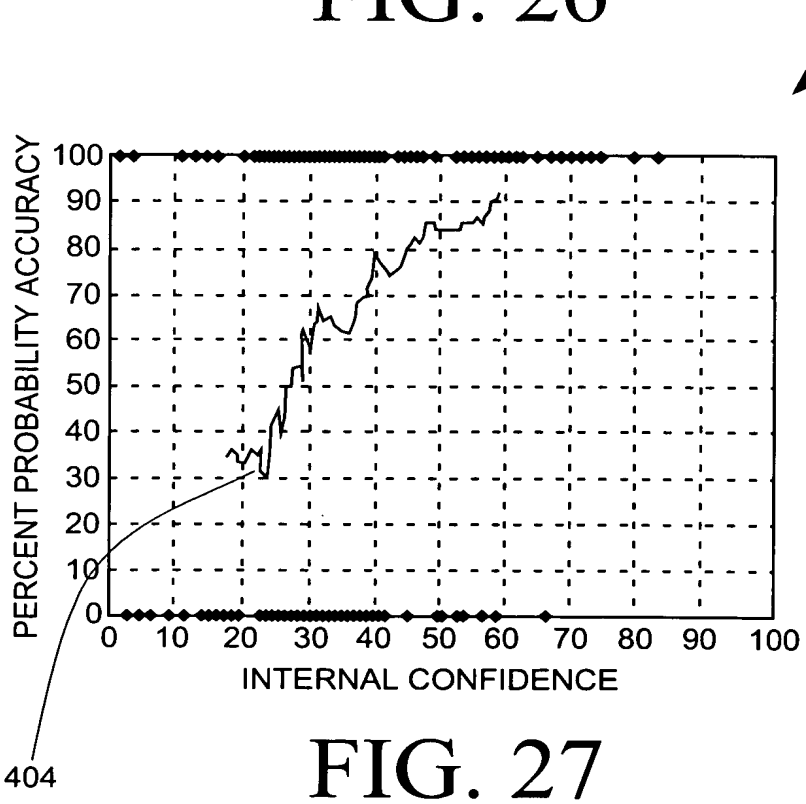

402
PERCENT PROBABILITY ACCURACY
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
INTERNAL CONFIDENCE
404

FIG. 27

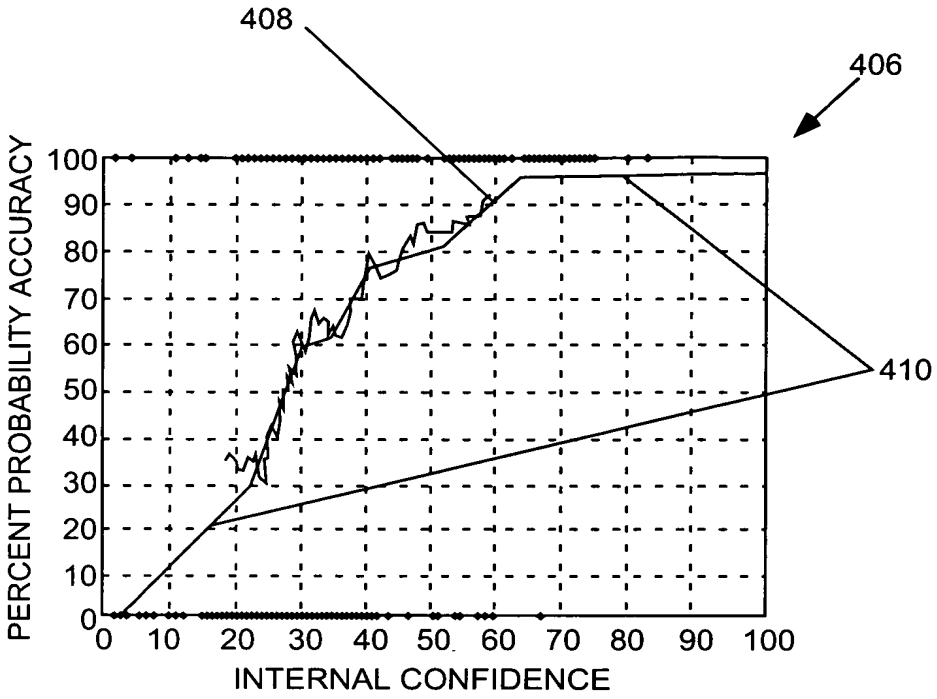

408
406
410
PERCENT PROBABILITY ACCURACY
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
INTERNAL CONFIDENCE

FIG. 28

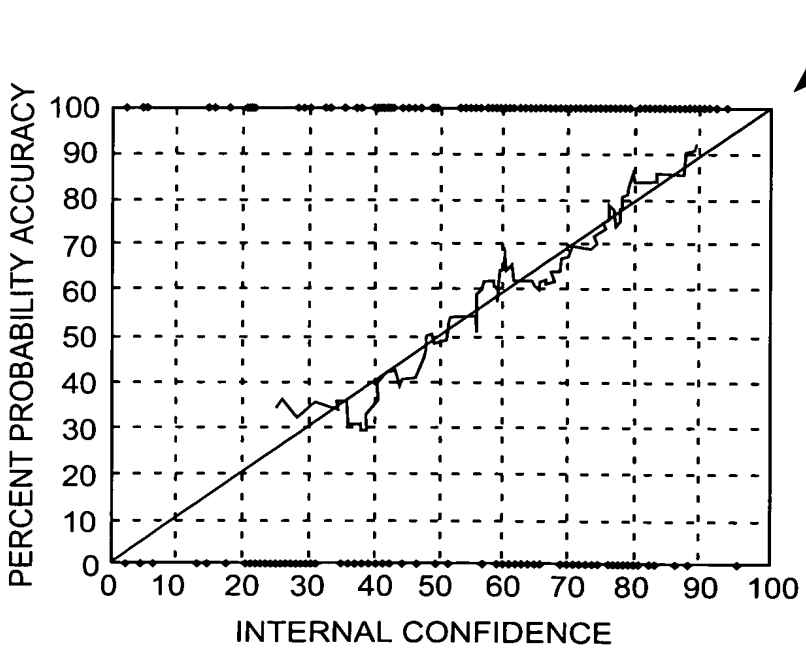

411
PERCENT PROBABILITY ACCURACY
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
INTERNAL CONFIDENCE

FIG. 29

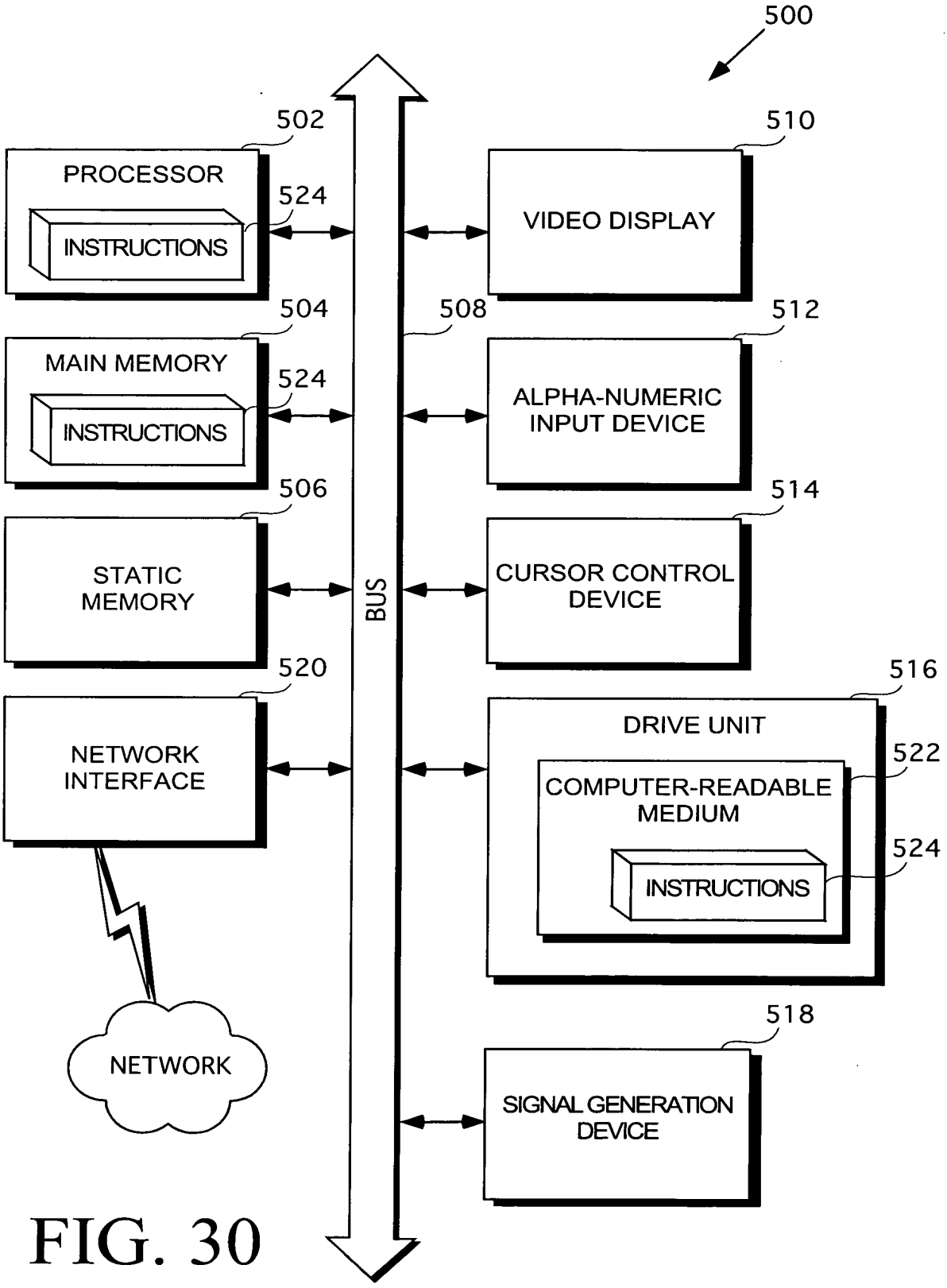

500
502
PROCESSOR
524
INSTRUCTIONS
504
MAIN MEMORY
524
INSTRUCTIONS
506
STATIC MEMORY
520
NETWORK INTERFACE
NETWORK
BUS
510
VIDEO DISPLAY
508
512
ALPHA-NUMERIC INPUT DEVICE
514
CURSOR CONTROL DEVICE
516
DRIVE UNIT
522
COMPUTER-READABLE MEDIUM
524
INSTRUCTIONS
518
SIGNAL GENERATION DEVICE

FIG. 30